

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
15 August 2002 (15.08.2002)

PCT

(10) International Publication Number  
**WO 02/062379 A2**

(51) International Patent Classification<sup>7</sup>: **A61K 39/08** (74) Agent: O'BRIEN JOHN A AND WELDON, Michael J; c/o John A. O'Brien & Associates, Third Floor, Duncairn House, 14 Carysfort Avenue, Blackrock, County Dublin (IE).

(21) International Application Number: PCT/IE02/00017

(22) International Filing Date: 11 February 2002 (11.02.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
2001/0137 9 February 2001 (09.02.2001) IE

(71) Applicant (*for all designated States except US*): **THE PROVOST, FELLOWS AND SCHOLARS OF THE COLLEGE OF THE HOLY AND UNIDIVIDED TRINITY OF QUEEN ELIZABETH [IE/IE]**; Near Dublin, College Green, Dublin 2 (IE).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **DOYLE, Rachael [IE/IE]**; 19 Deerpark Avenue, Castleknock, Dublin 15 (IE). **KELLEHER, Dermot [IE/IE]**; 30 Royal Terrace West, Dun Laoghaire, County Dublin (IE). **WINDLE, Henry J. [IE/IE]**; 15 Cherryfield Avenue Upper, Ranelagh, Dublin 6 (IE). **WALSH, James, Bernard [IE/IE]**; 3 Ardluig Park, Blackrock, County Dublin (IE). **DEIRDRE, Ni, Eidhin [IE/IE]**; 15 Watkins Buildings, The Coombe, Dublin 8 (IE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, IS, IT, FR, GB, GR, IE, PT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: CLOSTRIDIUM DIFFICILE VACCINE

(57) Abstract: A vaccine for the treatment or prophylaxis of *C. difficile* associated disease comprises a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans. The gene encodes a *C. difficile* surface layer protein, StpA or variant or homologue thereof. The peptide/polypeptide is a *C. difficile* surface layer protein, StpA or variant or homologue thereof. The vaccine may comprise a chimeric nucleic acid sequence.

**WO 02/062379 A2**

**"*Clostridium difficile* vaccine"**

**Introduction**

5       The invention relates to vaccines to provide immunological protection against *C. difficile* infection.

**Background**

10      *Clostridium difficile* is a common nosocomial pathogen and a major cause of morbidity and mortality among hospitalised patients throughout the world [Kelly et al., 1994]. Outbreaks of *C. difficile* have necessitated ward and partial hospital closure. With the increasing elderly population and the changing demographics of the population, *C. difficile* is set to become a major problem in the 21st century. The 15 spectrum of *C. difficile* diseases range from asymptomatic carriage to mild diarrhoea to fulminant pseudomembranous colitis. Host factors rather than bacterial factors appear to determine the response to *C. difficile* [Cheng et al., 1997; McFarland et al., 1991; Shim et al., 1998].

20      Reports indicate that hypogammaglobulinaemia in children appears to predispose to the development of disease due to *C. difficile* and that therapy with intravenously administered gamma globulin can be associated with the clinical resolution of chronic relapsing colitis due to *C. difficile* disease [Leung et al., 1991; Pelmutter et al., 1985]. A study by Mulligan et al. [1993] found elevated levels of 25 immunoglobulins reactive with *C. difficile* in asymptomatic carriers as opposed to symptomatic patients. Recently it has been shown that patients who became colonised with *C. difficile* who had relatively low levels of serum IgG antibody against toxin A had a much greater risk of developing *C. difficile* diarrhoea [Kyne et al., 2000].

30      It is clear that any advance in the understanding of *C. difficile* disease and methods of preventing or treating *C. difficile* diarrhoea (CDD) and other related diseases will be of major therapeutic potential.

Statements of Invention

According to the invention there is provided a vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans.

The invention also provides a vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising a *C. difficile* gene or *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof to which immunoreactivity is detected in individuals who have recovered from *C. difficile* infection.

Preferably the gene encodes a *C. difficile* surface layer protein, SlpA or variant or homologue thereof.

Preferably the peptide/polypeptide is a *C. difficile* surface layer protein, SlpA or variant or homologue thereof.

Most preferably the vaccine comprises a chimeric nucleic acid sequence. Preferably the chimeric nucleic acid sequence is derived from the 5' end of the gene, encoding the mature N-terminal moiety of SlpA from *C. difficile*.

In one embodiment of the invention the vaccine comprises a chimeric peptide/polypeptide. Preferably the amino acid sequence of the chimeric peptide/polypeptide is derived from the mature N-terminal moiety of SlpA from *C. difficile*.

Preferably the vaccine of the invention contains an amino acid sequence SEQ ID No.1 or a derivative or fragment or mutant or variant thereof.

Preferably the vaccine contains an amino acid sequence SEQ ID No.2 or a derivative or fragment or mutant or variant thereof.

In one embodiment of the invention the vaccine contains a nucleotide sequence SEQ ID No.3 or a derivative or fragment or mutant or variant thereof; a nucleotide sequence SEQ ID No.4 or a derivative or fragment or mutant or variant thereof; a nucleotide sequence SEQ ID No.5 or a derivative or fragment or mutant or variant thereof; a nucleotide sequence SEQ ID No.6 or a derivative or fragment or mutant or variant thereof; a nucleotide sequence SEQ ID No.7 or a derivative or fragment or mutant or variant thereof; a nucleotide sequence SEQ ID No.8 or a derivative or fragment or mutant or variant thereof; a nucleotide sequence SEQ ID No.9 or a derivative or fragment or mutant or variant thereof or a nucleotide sequence SEQ ID No.10 or a derivative or fragment or mutant or variant thereof.

Preferably the vaccine of the invention is in combination with at least one other *C. difficile* sub-unit.

The invention provides a vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising the mature N-terminal moiety of a surface layer protein, SlpA of *C. difficile* or variant or homologue thereof which is immunogenic in humans.

Most preferably the N-terminal moiety of SlpA contains an amino acid sequence SEQ ID No. 1.

In one embodiment of the invention the N-terminal moiety of SlpA contains an amino acid sequence SEQ ID No. 2.

The invention also provides a vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising an immunodominant epitope derived

from a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans.

5 Preferably the vaccine of the invention comprises a pharmaceutically acceptable carrier. Most preferably the vaccine is in combination with a pharmacologically suitable adjuvant. Ideally the adjuvant is interleukin 12. Alternatively the adjuvant may be a heat shock protein.

10 In one embodiment of the invention the vaccine comprises at least one other pharmaceutical product.

The pharmaceutical product may be an antibiotic, selected from one or more metronidazole, amoxycillin, tetracycline or erythromycin, clarithromycin or tinidazole.

15 15 In one embodiment of the invention the pharmaceutical product comprises an acid-suppressing agent such as omeprazole or bismuth salts.

20 The vaccine of the invention may be in a form for oral administration, intranasal administration, intravenous administration or intramuscular administration.

In one embodiment of the invention the vaccine includes a peptide delivery system.

25 The invention also provides an immunodominant epitope derived from a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof. Preferably the *C. difficile* peptide/polypeptide contains an amino acid sequence SEQ ID No.1 or SEQ ID No.2 or a derivative or fragment or mutant or variant thereof.

30 30 In one embodiment of the invention the *C. difficile* peptide/polypeptide contains an amino acid sequence SEQ ID No.3 or SEQ ID No.4 or SEQ ID No.5 or SEQ ID No.6 or SEQ ID No.7 or SEQ ID No.8 or SEQ ID No. 9 or SEQ ID No. 10 or a derivative or fragment or mutant or variant thereof.

The invention further provides a chimeric nucleic acid sequence derived from the 5' end of the *slpA* gene encoding the mature N-terminal moiety of *SlpA* from *C. difficile* which is immunogenic in humans.

5 The invention also provides a chimeric peptide/polypeptide wherein the amino acid sequence of the chimeric peptide/polypeptide is derived from the mature N-terminal moiety of *SlpA* from *C. difficile*.

10 The invention provides a *C. difficile* peptide comprising SEQ ID No. 1 or SEQ ID No. 2 or SEQ ID No. 3 or SEQ ID No. 4 or SEQ ID No. 5 or SEQ ID No. 6 or SEQ ID No. 7 or SEQ ID No. 8 or SEQ ID No. 9 or SEQ ID No. 10.

15 One aspect of the invention provides for the use of a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans in the preparation of a medicament for use in a method for the treatment or prophylaxis of *C. difficile* infection or *C. difficile* associated disease in a host.

Preferably the medicament which is prepared is a vaccine of the invention.

20 The invention also provides a method for preparing a vaccine for prophylaxis or treatment of *C. difficile* associated disease, the method comprising;

25 obtaining a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans; and

30 forming a vaccine preparation comprised of said gene or peptide/polypeptide or derivative or fragment or mutant or variant, which is suitable for administration to a host and which when administered raises an immune response.

Preferably the *C. difficile* peptide/polypeptide contains an amino acid sequence SEQ ID No.1 or SEQ ID No.2 or a derivative or fragment or mutant or variant thereof.

5       Most preferably the *C. difficile* gene contains an amino acid sequence SEQ ID No.3 or SEQ ID No.4 or SEQ ID No.5 or SEQ ID No.6 or SEQ ID No.7 or SEQ ID No.8 or SEQ ID No.9 or SEQ ID No.10 or a derivative or fragment or mutant or variant thereof.

10      The invention further provides a method for prophylaxis or treatment of *C. difficile* associated disease, the method comprising;

15      obtaining a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans;

20      forming a vaccine preparation comprised of said gene or peptide/polypeptide or derivative or fragment or mutant or variant, and

25      administering the vaccine preparation to a host to raise an immune response.

30      One aspect of the invention provides monoclonal or polyclonal antibodies or fragments thereof, to a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans.

35      Another aspect of the invention provides monoclonal or polyclonal antibodies or fragments thereof, to *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof to which immunoreactivity is detected in individuals who have recovered from *C. difficile* infection.

40      The invention also provides purified antibodies or serum obtained by immunisation of an animal with a vaccine of the invention.

The invention provides the use of the antibodies or fragments of the invention in the preparation of a medicament for treatment or prophylaxis of *C. difficile* infection or *C. difficile* associated disease.

5 Preferably the antibodies or serum are used in the preparation of a medicament for treatment or prophylaxis of *C. difficile* infection or *C. difficile* associated disease.

Most preferably the antibodies or fragments or serum of the invention are used in passive immunotherapy for established *C. difficile* infection.

10 In one embodiment of the invention the antibodies or fragment or serum of the invention are used for the eradication of *C. difficile* associated disease.

15 The invention also provides use of interleukin 12 as an adjuvant in *C. difficile* vaccine.

The invention further provides use of humanised antibodies or serum for passive vaccination of an individual with *C. difficile* infection.

20 Brief Description of the Drawings

25 The invention will be more clearly understood from the following description thereof given by way of example only with reference to the accompanying figures, in which:-

30 Fig. 1A is a Western blot showing recognition of antigens from a crude extract of *C. difficile* 171500 (PCR type 1) by serum antibodies from a patient infected with this strain. Lane 1: Pre-infection; Lane 2: Early acute; Lane 3: Late acute; Lane 4: Convalescent;

Fig. 1B is a Western blot showing recognition of antigens from a crude extract of *C. difficile* 170324 (PCR type 12) by serum antibodies from a patient infected with this strain. Lane 1: Pre-infection; Lanes 2-5: Acute; Lanes 6-7: Convalescent;

5

Fig. 2. is a Western blot showing recognition of antigens from two *C. difficile* strains of different type by serum from convalescent patients.

10

- Lane 1: Strain 170324 (PCR type 12), crude antigen preparation
  - Lane 2: Strain 170324, surface layer protein preparation
  - Lane 3: Strain 171500 (PCR type 1), crude antigen preparation
  - Lane 4: Strain 171500, surface layer protein preparation.
- Molecular mass markers (kDa) are shown on the left; and

15

Fig. 3 is an SDS-PAGE gel showing crude SLP preparations from selected strains of *C. difficile*. The gel contains 12% acrylamide, and has been stained for protein with Coomassie Blue. Each lane contains 5 µg of protein. Molecular weight markers are shown on the left.

20

- Lane 1: 171500 (PCR type 1)
- Lane 2: 172450 (PCR type 5)
- Lane 3: 170324 (PCR type 12)
- Lane 4: 171448 (PCR type 12)
- Lane 5: 171862 (PCR type 17)
- Lane 6: 173644 (PCR type 31)
- Lane 7: 170444 (PCR type 46)
- Lane 8: 170426 (PCR type 92)

25

#### Detailed Description of the invention

30

Two antigenic peptides containing SEQ ID No. 1 and SEQ ID No. 2, associated with two common infecting types of *C. difficile*, were found to be immunogenic in humans. The antigenic peptides were found to induce a strong immune response in

individuals who recover from *C. difficile* infection. Individuals who have recovered from *C. difficile* infection are those individuals who have been exposed to *C. difficile* or something strongly related and have recovered. This includes individuals where a carrier state exists in that the *C. difficile* infection has not and will not necessarily become clinically significant.

These antigenic peptides were found to be products of the *slpA* gene from *C. difficile* which is the structural gene for the surface layer protein, SlpA. The gene or its products are therefore ideal candidates for the preparation of vaccines against *C. difficile*.

Surface layer proteins (SLPs), also known as S-layers or crystalline surface layers, are associated with a wide range of bacterial species. They form a 2-dimensional array, which covers the surface of the cell completely, and grows with the cell [Sleytr et al., 1993]. The molecular weight can range from 40 000 to 200 000 Da. The proteins are typically acidic, contain a large proportion of hydrophobic amino acid residues, and have few or no sulphur-containing amino acid residues. Glycosylated S-layer proteins occur in some species. The precise function of S-layers is not always known, but since they comprise approximately 15% of the cell protein, it seems likely that they are important for *in vivo* functioning of the organism. In Gram positive organisms, the SLP has been shown to delay or prevent the excretion of degradative enzymes from the cell to the outside milieu, and may thereby create a space analogous to the periplasmic space of Gram negative bacteria. Many pathogenic species possess SLPs, which have been ascribed functions such as antiphagocytosis (*Campylobacter fetus*), and inhibition of complement-mediated killing (*Aeromonas salmonicida*).

Kawata et al. [1984] described the SLPs of *Clostridium difficile*. They showed the S-layer to be composed of 2 polypeptides, and demonstrated size heterogeneity for the polypeptides from different strains. Delmée et al. [1986] showed that crude extracts from *C. difficile* strains of different serotype showed different polypeptide profiles in SDS-PAGE. Poxton et al. [1999] made similar observations using purified SLP preparations. Slide agglutination [Delmée et al., 1990] has identified 21 different serotypes, apparently distinguished by the heterogeneity of the SLP.

Pantosti et al. [1989] isolated *C. difficile* from a number of patients with antibiotic-associated diarrhoea, and prepared SLPs from them.. Cerquetti et al. [2000] published N-terminal sequences of SLPs from several strains, indicating wide differences between strains.. In 2000 the complete DNA sequence of the *C. difficile* genome was published (available at web address [http://www.sanger.ac.uk/Projects/C\\_difficile/](http://www.sanger.ac.uk/Projects/C_difficile/)).

The peptides of the invention were found to be encoded by a single open reading frame (ORF) named *slpA* from *C. difficile*. The peptides identified in our clinical study correspond to a lower molecular weight moiety of the *slpA* gene product. Since an immune response is also mounted against a higher molecular weight *slpA* gene product (Fig. 2), this entity may also be included in a vaccine.

The *slpA* gene has been sequenced from a number of strains corresponding to different PCR types. The sequences of strains 171500 (PCR type 1)(NCIMB 41081; PHLS R13537), 172450 (PCR type 5)(PHLS R12884), 170324 (PCR type 12) (NCIMB 41080; PHLS R12882),, 171448 (PCR type 12) (PHLS R13550), 171862 (PCR type 17) (PHLS R13702), 173644 (PCR type 31) (PHLS R13711), 170444 (PCR type 46) (PHLS R12883) and 170426 (PCR type 92) (PHLS R12871) with translations thereof are given in Appendices 1 to 8. Substantial variation in nucleotide and predicted amino acid sequence was found between strains of PCR types 1, 5, 12, 17 and 31. The genes from strains of PCR types 46 and 92 are almost identical in sequence to those of PCR type 12. When the DNA sequences of genes of different strains within a PCR type are compared, the sequences are almost if not quite identical, indicating that the potential for variation is not infinite. These findings are in agreement with serotyping studies [Delmée et al., 1986, 1990], and indicate that the production of an effective vaccine based on the *slpA* product is feasible. In this respect, the present invention includes all variant *slpA* genes and their products, individually and combined, fragments of them, and their mutants and derivatives.

One aspect of the invention provides the combination of immunodominant epitopes from the *slpA* gene products from various serotypes into a single vaccine. In this way a single vaccine may be used to immunise against several different *C. difficile* strains.

- The most common PCR types isolated from infections in the clinical study carried out at St. James's Hospital, Dublin, Ireland were PCR types 1 and 12. However, a vaccine which elicits an intense antibody response against many infecting types would be therapeutically very valuable. Recombinant DNA chimera, or several chimeras, encoding contiguous immunodominant epitopes may be made for use in the vaccine. The recombinant DNA may serve as the active component in a vaccine, or may be inserted into an appropriate expression system for the generation of a chimeric peptide vaccine in a suitable host.
- Chimeras can be generated by PCR amplification of the DNA encoding peptide regions of interest, incorporating cleavage sites for restriction endonucleases into the primers. The amplified fragments can thus be cleaved to generate compatible ends, and spliced together to create chimeras.
- The dominant epitopes may be identified by cleavage of the *sipA* products into fragments by agents which cleave at known sites, and by immunoblotting with homologous patient serum. Immunodominant peptides may be tested for their capacity to stimulate T-cell proliferative responses *in vitro*, using mouse splenic T-cells.
- DNA vaccination involves immunisation with recombinant DNA encoding the antigen or epitope of interest, cloned in a vector which promotes high level expression in mammalian cells. Typically, the vector is a plasmid vector which also replicates in a prokaryotic vector such as *Escherichia coli*, so that the DNA can be produced in quantity. Following immunisation, the plasmid enters a host cell, where it remains in the nucleus, and directs synthesis of the recombinant polypeptide. The polypeptide stimulates the production of neutralising antibodies, as well as activating cytotoxic T-cells.
- Using a DNA vaccine, it may be necessary to modify the DNA sequence to take account of codon usage in humans. The G+C content of mammalian DNA is much higher than that of *C. difficile*. The generation of such synthetic DNA molecules, essentially containing numerous silent mutations, is within the scope of the invention.

A peptide vaccine will ideally be made using recombinant peptides. Similar considerations apply as in the generation of a DNA vaccine with regard to expression in a different host, such as *Escherichia coli*, which has a different codon usage pattern to *C. difficile*. Problems of expression may be overcome by the use of a 5 special host strain which carries additional copies of rare tRNAs (e.g. *E. coli* BL21-CodonPlus<sup>TM</sup>-RIL from Stratagene), or by using *de novo* synthesis of a DNA segment carrying silent mutations which will enable normal expression in *E. coli*. There are many expression systems which are likely to allow high-level expression 10 of *slpA* genes in *E. coli*. An example is the pBAD/Thio TOPO vector of Invitrogen, in which expressed genes are under control of the arabinose promoter, which is subject to positive and negative control, enabling very tight control of expression. In this vector, the recombinant protein is typically fused to a modified thioredoxin 15 carrying several histidine residues which enable purification by nickel chromatography. The recombinant protein can be cleaved from the thioredoxin moiety by enterokinase enzyme.

Affinity chromatography may also be used with fixed antibodies or some other agent which strongly binds the peptide of interest to purify the protein from the native organism.

Purified immunogenic peptides may be used in combination with other *C. difficile* 20 sub-units as a combined vaccine against *C. difficile*. Potential candidates are the products of the other *slp* genes, which share limited homology with the *slpA* gene product and with the N-acetyl muramoyl L-alanine amidase, (CwlB), from *Bacillus subtilis*, and which may be involved in remodelling of the peptidoglycan.

Other purified proteins of *C. difficile* to which constitutive antibodies are detected 25 in individuals recovering from *C. difficile* infection are also within the scope of the present invention

A deposit of *Clostridium difficile* strain 171500, PCR type 1, was made at the 30 NCIMB on January 29, 2001, and accorded the accession number NCIMB 41081.

A deposit of *Clostridium difficile* strain 170324, PCR type 12, was made at the 35 NCIMB on January 29, 2001, and accorded the accession number NCIMB 41080.

Two peptides of the invention were found to contain the following sequences:

5           33kDa peptide

SEQ ID No. 1:       DKTKVETADQGYTVVQSKYK

10           31kDa peptide

SEQ ID No. 2       ATTGTQGYTVVKNDGKKAVK

15           The invention will be more clearly understood from the following examples.

20           Example 1. Clinical Study

25           Examination of sequential antibody responses to *C. difficile* among elderly patients who developed the disease was carried out. The study was based on the hypothesis that the host immune response influenced the development of *Clostridium difficile* disease. In particular we determined that a particular pattern of immune response to *C. difficile* antigens correlated with the outcome of CDD.

30           Materials and Methods

20           Patients

25           Serum was collected from over 300 patients and of these 30 patients developed CDD. The infecting strain (homologous strain) was grown from each patient. Strains of *C. difficile* were typed at the Anaerobe Reference Laboratory, Wales [O'Neill et al., 1996]. The most common strains isolated were PCR type 1 (n = 15) which is the most common type causing epidemics and PCR type 12 (n = 5) which is also a common hospital strain. Pre-infection serum samples were obtained from patients. Acute phase sera were then collected from patients who developed *C. difficile* disease. Convalescent sera were collected from patients who recovered.

30           Protein extracts of patients' infecting *C. difficile* strain were probed with the patients sera using Western blotting. IgG responses to the antigens were examined.

35           Western blotting

Proteins from SDS-PAGE gels were electroblotted (0.8mA/cm<sup>2</sup> for 1 h) to PVDF membrane using a semi-dry blotting apparatus (Atto). Primary antibodies (human

serum: 1/50 – 1/10,000 dilution) were detected using a 1/5000 dilution of anti-human IgG (horse radish peroxidase-conjugated) in combination with enhanced chemiluminescence (ECL). Blots were washed in phosphate buffered saline (pH 7.5) containing Tween 20 (0.1% v/v), and incubated in the same solution comprising dried skim milk (5% w/v) and antibodies at the appropriate concentration. Blots were exposed to Kodak X-OMAT film for various periods of time and developed.

### Results

Overall 5 patients made a full recovery and new antibody responses to previously unrecognised antigens were evident in 4 of these patients. Three of these patients had *C. difficile* belonging to PCR type 1 and one patient had *C. difficile* PCR type 12. These patients developed an acute phase antibody response to previously unrecognised *C. difficile* antigens which persisted during convalescence (Figs. 1A and 1B). These antigens were recognised by antibodies from patients who recovered and represent potential candidate vaccine antigens. Fig 1A shows a strong reaction of convalescent antibodies was observed with the 33 kDa antigen (Lane 4, arrow). Fig 1B shows a strong reaction of convalescent antibodies was observed with the 31 kDa antigen (Lanes 6 and 7, arrow).

These antibody responses have also been found in some controls in the same ward who were also on antibiotics but who did not develop CDD.

### Example 2. Further characterisation of protective antigens

#### Materials and Methods

Partial purification and N-terminal sequencing of the 33 kDa and the 31 kDa proteins The antigens were partially purified from *C. difficile* based on their molecular weight using preparative continuous-elution SDS-PAGE on a model 491 Prep-Cell (Bio-Rad). The appropriate antigens were subsequently identified on Western blots probed with serum obtained from individuals who recovered from *C. difficile* infection.

#### Preparation of surface layer proteins (SLPs)

SLPs were purified from *C. difficile* by extracting washed cells with 8 M urea, in 50 mM Tris HCl, pH 8.3 in the presence of a cocktail of protease inhibitors

(Complete®, Boehringer Mannheim), for 1 h at 37°C, followed by centrifugation for 19 000 x g for 30 min. The SLPs were recovered in the supernatant and dialysed to remove the urea [Cerquetti et al., 2000].

5      Results

The immunodominant protein which was associated with a positive outcome from *C. difficile* strain 171500 (PCR type 1) was identified and purified using preparative SDS-PAGE. The N-terminal region of the protein was sequenced using an Applied Biosystems Procise Sequencer, viz DDKTKVETADQGYTVVQSKYK (SEQ ID No. 1)

10     The antigen which was associated with a protective antibody response from the *C. difficile* strain 170324 (PCR type 12) was identified and the N-terminal sequence obtained, viz ATTGTQGYTVVKNDGKKAVK (SEQ ID No. 2).

15     These sequences were used to interrogate the *C. difficile* genome sequence using the TBLASTN programme, which compared our query sequences with those of the genome project (available at web address [http://www.sanger.ac.uk/Projects/C\\_difficile/](http://www.sanger.ac.uk/Projects/C_difficile/)), translated in all 6 possible reading

20     frames. A nearly identical stretch of sequence was identified when the sequence from strain 170324 (type 12) was used for interrogation. The same stretch of sequence was picked up with the sequence from strain 171500 (type 1) was used, although the identity was much less strong. Since the homologous sequence belonged to an open reading frame encoding a 719-residue peptide, this result was somewhat surprising. However, when the N-terminal sequences from the higher molecular weight SLP component were later published by Cerquetti et al [2000], it became apparent that they were encoded downstream along the same gene, subsequently identified as *sdpA*, and the reason for the discrepancy in size between the gene and its products became readily apparent.

25     30     The purified SLPs from strains 171500 (PCR type 1) and 170324 (PCR type 12) showed strong reactivity with homologous convalescent serum, and co-migrated with the dominant antigens detected in crude cell extracts as shown in Fig. 2. Lanes 1 and 3 contain crude antigen preparations from PCR types 1 and 12 respectively, and Lanes 2 and 4 contain SLP preparations from PCR types 1 and 12, respectively.

Panel A was probed with serum from a patient recovering from infection with PCR type 1, and Panel B was probed with serum from a patient recovering from infection with PCR type 12. Each serum detected 2 major antigens in the infecting strain (Panel A, Lane 3); (Panel B, Lane 1), which co-migrated with the 2 SLPs (Panel A, 5 Lane 4; Panel B, Lane 2), with which the sera also reacted strongly. Note that serum from the patient infected with the PCR type 1 strain recognised the higher molecular weight SLP from the PCR type 12 strain (Panel A, Lanes 1 and 2), whereas the converse did not occur (Panel B, Lanes 3 and 4). There is no apparent antigenic cross-reactivity with regard to the lower molecular weight SLPs.

10

SLPs were prepared from selected strains by urea extraction, and subjected to SDS-PAGE and staining with Coomassie Blue (Fig. 3). Most strains showed a characteristic profile, with two major bands located in the 29 000 to 36 000 and 45 000 to 50 000 molecular weight range. An exception was strain 172450 (Fig. 3, 15 Lane 2), which showed a single, high molecular weight band, approximately 43 000 in size.

#### Cloning, sequencing and analysis of *sipA* genes

The nucleotide sequences of the *sipA* genes from the two sample strains of *C. difficile* (PCR types 1 and 12, deposited at the NCIMB) and of several others (PCR types 5, 12, 17, 31, 46 and 92, available from the Anaerobe Reference Unit at the Department of Medical Microbiology and Public Health Laboratory, Cardiff, Wales were obtained. The *sipA* gene and flanking sequence was amplified by polymerase chain reaction from genomic DNA prepared from *C. difficile* using a commercial kit (Puregene® DNA isolation kit for yeast and Gram positive bacteria, Gentra systems Minneapolis, MN). The forward primer (5' ATGGATTATTATAGAGATGTGAG 3'), was based on sequence from the genome sequencing project, starting 112 nucleotides upstream from the start of the *sipA* open reading frame. Two reverse primers were used, depending on the PCR type. A downstream primer (5' CTATTAAAGTTTATTAAAACCTTATATTAC 3') was used to amplify *sipA* 25 from PCR types 12, 17, 31, 46 and 92. A reverse primer based on the 3' end of the *sipA* open reading frame from strain 630 and the subsequent nonsense codon (5' TTACATATCTAATAAATCTTCATTTGTTATAACTG 3') was used to 30

amplify *slpA* from PCR types 1 and 5. The choice of primer for the latter two PCR types may have resulted in a small number of systematic errors in the nucleotide sequence obtained. PCR was carried out using HotStar™ Taq polymerase (Qiagen Ltd., Crawley, West Sussex, UK) according to the manufacturer's instructions. A single fragment of approximately 2 kb was obtained for each strain, which was then cloned into the pBAD/Thio TOPO vector (Invitrogen, Groningen, Netherlands). Inserts were sequenced from both ends by standard procedures in commercial facilities at MWG (Wolverton Mill South, Milton Keynes, UK) and Cambridge University. New primers were designed on the basis of initial sequencing results, enabling sequencing of both strands to be completed (a process known as chromosome walking).

The results are shown in Appendices 1-8.

The nucleotide sequences were translated to enable prediction of the amino acid sequence(s) of the product(s) (Appendices 1-8). The N-terminal sequences obtained experimentally for the low molecular weight protective antigens from strains 171500 (PCR type 1) and 170324 (PCR type 12) were almost identical to those predicted from the nucleotide sequences of their respective *slpA* genes (18/20 identical residues for strain 171500, and 19/20 identical residues for strain 170324).

Appendix 1 shows the open reading frame with translation for *slpA* from strain 171500 (PCR type 1), SEQ ID No 3. Since the reverse primer was based on the 35 nucleotides from the 3' end of the *slpA* gene, the sequence is not necessarily 100% accurate in this region. However, this part of the gene does not seem to vary greatly from strain to strain.

Appendix 2 shows the open reading frame with translation for *slpA* from strain 172450 (PCR type 5), SEQ ID No 4. Again, the sequence obtained for the 3' 35 nucleotides is not fully reliable. This gene is considerably smaller than the other *slpA* genes sequenced, and shows strong sequence divergence from the other PCR types examined.

Appendix 3 shows the open reading frame with translation for *slpA* from strain 170324 (PCR type 12), SEQ ID No 5. This gene showed a single base difference

when compared with the strain used for the genome sequencing project, strain 630, of the same PCR type. The deduced amino acid sequence is identical.

5 Appendix 4 shows the open reading frame with translation for *slpA* from strain 171448 (PCR type 12), SEQ ID No 6. This gene was almost identical in sequence to that from strain 170324.

Appendix 5 shows the open reading frame with translation for *slpA* from strain 171862 (PCR type 17), SEQ ID No 7.

10 Appendix 6 shows the open reading frame with translation for *slpA* from strain 173644 (PCR type 31), SEQ ID No 8. Like the *slpA* from strain 172450, this sequence is very dissimilar to those of *slpA* genes from other PCR types encountered.

15 Appendix 7 shows the open reading frame with translation for *slpA* from strain 170444 (PCR type 46), SEQ ID No 9. This sequence is virtually identical to that obtained for *slpA* from PCR type 12 and 92 strains.

20 Appendix 8 shows the open reading frame with translation for *slpA* from strain 170426 (PCR type 92), SEQ ID No 10. This sequence is virtually identical to that obtained for *slpA* from PCR type 12 and 46.

25 The cleavage site of the putative signal sequences from both genes was determined from experimental evidence (the N-terminal sequence of the mature proteins as determined by Edman degradation), and by the prediction tool of the Centre for Biological Sequence Analysis at the Technical University of Denmark [Nielsen et al., 1997]. The site for cleavage of the *slpA* gene product to form the mature SLPs was predicted from experimental [Cerquetti et al., 2000, Karjalainen et al., 2001 and Calabi et al., 2001]. The cleavage site is typically preceded by the motif TKS. However, the relevant motif is likely to be TKG in strain 173644 (PCR type 31). No obvious motif appeared for strain 172450 (PCR type 5). However, the protein produced by type 5 strains does appear to be cleaved; hence we predicted the site to

occur at a point where the SLP sequence aligns with the cleavage sites of other PCR types.

5       The molecular weight and isoelectric point was calculated for each of the predicted mature proteins by the ExPASy server of the Swiss Institute for Bioinformatics (Table 1). In general, the calculated molecular weights were in fair agreement with apparent molecular masses determined from migration in gels (Fig. 3). No lower molecular weight band was apparent for Strain 172450 (PCR type 5; Lane 2). However, a higher molecular weight band is present, which is similar in size to the 10 predicted weight for the C-terminal moiety. We observed a similar profile for another type 5 strain. It is possible that the lower molecular weight species is subject to degradation in this strain. Another possibility is that it is heavily glycosylated, which can affect staining. All peptides had a predicted isoelectric point below 7, typical of acidic proteins, and characteristic of SLPs in general [Sleyter et al, 1993].

15

Table 1

<i>C. difficile</i> strain (PCR type)	pI (N-terminal)	pI (C-terminal)	MW (N-terminal)	MW (C-terminal)
171500 (Type 1)	4.83	4.66	33365.41	44220.37
172450 (Type 5)	4.86	4.65	19364.46	42757.63
170324 (Type 12)	4.92	4.58	34228.25	39522.24
171448 (Type 12)	4.98	4.58	34156.18	39492.21
171862 (Type 17)	5.09	4.53	33783.73	39407.11
173644 (Type 31)	5.05	4.56	33626.48	41821.69
170444 (Type 46)	5.06	4.58	34230.31	39522.24
170426 (Type 92)	4.99	4.58	34242.32	39522.24

20       The translated nucleotide sequences were compared with published SlpA sequences (EMBL Accession numbers AJ300676, and AJ300677 for examples from PCR types 1, and 17 respectively; strain 630 available from the Sanger Institute for PCR type 12; EMBL Accession number AY004256 for a variant from an unnamed PCR type). The Clustal W alignment programme, which is freely available, was used. Where 25 SlpA sequences from our isolates were compared with those of other strains of the same PCR types, they were found to be nearly or quite identical. This observation

indicates, together with existing knowledge from serotyping, that the number of variants of *sfpA* is not infinite, and that natural evolution of the gene is not rapid. Table 2 shows a compilation of homologies, based on amino acid residue identity, for the different translated sequences measured against published sequences.

5 Homologies are compiled for the predicted mature peptides, either combined (Table 2A) or as N-terminal (low molecular weight, less conserved moiety) (Table 2B) and C-terminal (high molecular weight, more conserved) (Table 2C) mature peptides according to predicted cleavage sites. It is clear that the SfpA sequences from strains 172450 (PCR type 5) and 173644 (PCR type 31) are quite distinct particularly with respect to N-terminal region.

10

Table 2A

Strain.type	630 (type 12)	AJ300676 (type 1)	AJ300677 (type 17)	AY004256 (type unknown)
171500.type1	55.2	99.7	55.4	56.42
172450.type5	49.8	54.0	49.9	47.77
170324.type12	100.0	57.8	81.7	59.77
171448.type12	99.7			
171862.type17	82.3	58.7	100	57.54
173644.type31	57.9	59.2	60.1	56.88
170444.type46	99.6			
170426.type92	99.9			

15

Table 2B

Strain.type	630 (type 12)	AJ300676 (type 1)	AJ300677 (type 17)	AY004256 (type unknown)
171500.type1	35.4	100	34.5	33.54
172450.type5	31.6	32.2	31.0	24.58
170324.type12	100	34.9	64.6	36.14
171448.type12	99.7			
171862.type17	64.3	34.4	100	31.55
173644.type31	37.5	34.1	41.3	31.86
170444.type46	99.1			
170426.type92	99.7			

Table 2C

Strain.type	630 (type 12)	AJ300676 (type 1)	AJ300677 (type 17)	AY004256 (type unknown)
171500.type1	70.2	99.5	71.2	73.80
172450.type5	58.4	60.4	63.0	57.60
170324.type12	100	77.3	97.1	80.00
171448.type12	99.7			
171862.type17	97.3	78.8	100	79.62
173644.type31	74.1	78.9	75.1	75.38
170444.type46	100			
170426.type92	100			

5

The term antibody used throughout the specification includes but is not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by a Fab expression library.

10

The antibodies and fragments thereof may be humanised antibodies. Neutralising antibodies such as those which inhibit biological activity of the substance amino acid sequence are especially preferred for diagnostics and therapeutics.

15

Antibodies both polyclonal and monoclonal which are directed against epitopes obtainable from a polypeptide or peptide of the present invention are particularly useful in diagnosis and those which are neutralising are useful in passive immunotherapy.

20

Antibodies may be produced by any of the standard techniques well known in the art.

25

A therapeutically effective amount of the polypeptide, polynucleotide, peptide or antibody of the invention in the form of pharmaceutical composition may be administered. The composition may optionally comprise a pharmaceutically acceptable carrier, diluent or excipients and including combinations thereof. The pharmaceutical composition may be used in conjugation with one or more additional pharmaceutically active compounds and/or adjuvants.

Different adjuvants depending on the host may be used to increase immunological response. The adjuvant may be selected from the group comprising Freunds, mineral gels such as aluminium hydroxide and surface active substances.

- 5      The vaccine of the invention may be in the form of an immune modulating composition or pharmaceutical composition and may be administered by a number of different routes such as by injection (which includes parenteral, subcutaneous and intramuscular injection) intranasal, intramuscular, mucosal, oral, intra-vaginal, urethral or ocular administration. There may be different formulation/composition  
10     requirements dependent on the different delivery systems.

- 15     The invention is not limited to the embodiments hereinbefore described which may be varied in detail.

References

- Calabi E., Ward S., Wren B., Paxton T., Panico M., Morris H., Dell A., Dougan G., Fairweather N. (2001). Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Mol. Microbiol.* 40:1187-1199.
- 5 Cerquetti M., Molinari A., Sebastianelli A., Diociaiuti M., Petruzzelli R., Cap C., Mastrandriano P. (2000). Characterization of surface layer proteins from different *Clostridium difficile* clinical isolates. *Microbial Pathogenesis*, 28:363-372.
- Cheng S.H, Lu J.J, Young T.G, Perng C.L, Chi W.M. (1997) *Clostridium difficile*-associated 10 diseases: comparison of symptomatic infection versus carriage on the basis of risk factors, toxin production, and genotyping results. *Clin Infect Dis* ; 25: 157-8.
- Delmée M., Laroche Y., Avesani V., Cornelis G. (1986). Comparison of serogrouping and polyacrylamide gel electrophoresis for typing *Clostridium difficile*. *J. Clin. Microbiol.* 24:991-994.
- 15 Delmée M., Avesani V., Delferrière N., Burtonboy G. (1990). Characterization of flagella of *Clostridium difficile* and their role in serogrouping reactions.
- Karjalainen T., Waligora-Dupriet A.-J., Cerquetti M., Spigaglia P., Maggioni A., Mauri P., Mastrandriano P., (2001). Molecular and genomic analysis of genes encoding surface-anchored proteins from *Clostridium difficile*. *Infect. Immun.* 69:3442-3446.
- 20 Kawata T., Takeoka A., Takumi K., Masuda K. (1984). Demonstration and preliminary characterization of a regular array in the cell wall of *Clostridium difficile*. *FEMS Microbiol. Lett* 24:323-328.
- Kelly, C.P., Pothoulakis C and LaMont J.T. *Clostridium difficile* colitis. *New England Journal of Medicine*. 1994 330: 257-262.
- 25 Kyne L, Warny M, Qamar A, Kelly C. Asymptomatic carriage of *Clostridium difficile* and serum levels of IgG antibody against Toxin A. *New England Journal of Medicine* 2000; 390:7.
- Leung Y.M, Kelly C.P, Boguniewicz M, Pothoulakis C, LaMont J.T, Flores A. Treatment with intravenous gamma globulin of chronic relapsing colitis by *Clostridium difficile*; toxin: *J. Pediatr* 1991; 118: 633-7.
- 30 McFarland L.V, Elmer G.W, Stamm W.E, Mulligan M.E. Correlation of immunoblot type, enterotoxin production, and cytokine production with clinical manifestation of *Clostridium difficile* infection in a cohort of hospitalised patients. *Infect Immun.* 1991; 59: 2456-62.

- Mulligan M.E, Miller S.D, McFarland L.V, Fung H.C, Kwok R.Y. Elevated levels of serum immunoglobulins in asymptomatic carriers of *Clostridium difficile*. *Clin Infect Dis* 1993; 16(Suppl 4); S239-44.
- 5 Nielsen H., Engelbrecht J., Brunak S., von Heijne G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10:1-6.
- O'Neill G.L., Ogunsota F.T., Brazier J.S., Duerdon B.I. Modification of a PCR ribotyping method for application as a routine typing scheme for *Clostridium difficile*. *Anaerobe* (1996) 2, 205-209.
- 10 Pantosti A, Cerquetti M, Viti F, Ortisi G, Mastrantonio P. Immunoblot Of Serum Immunoglobulin G Response to Surface Proteins of *Clostridium difficile* in Patients With Antibiotic Associated Diarrhoea. *J. Clin Microbiol* 1989; 27; 2594-7.
- Pelhamer D.H, Leichtn A.M, Goldman H, Winter H.S. Chronic diarrhoea associated with 15 hypogammaglobulinaemia and enteropathy in infants and children: *Dig Dis Sci* 1985; 30; 1149-55.
- Poxton I.R., Higgins P.G., Currie C.G., McCoubrey J. (1999). Variation in the cell surface proteins of *Clostridium difficile*. *Anaerobe* 5:213-215.
- Shim J. Johnson S, Samone M, Bliss DZ, Gerding D.N. Primary symptomless colonisation by *Clostridium difficile* and decreased risk of subsequent diarrhoea. *The Lancet* Vol 20 351 1998: 633-5.
- Sleytr U.B., Messner P., Pum D., Sára M. (1993). Crystalline bacterial cell surface layers. *Mol. Microbiol.* 10:911-916.

## Appendix 1

SEQ ID No. 3. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 171500, PCR type 1, with translation. The putative secretory signal cleavage site ( $\Delta$ ) and site of cleavage to form the two mature SLPs ( $\blacklozenge$ ) are indicated.

5 1 ATGAATAAGAAAAATATAGCAATAGCTATGTCAGGTTAACAGTTTAGCTCGGCTGCA 60

10 1 M N K K N I A I A M S G L T V L A S A A  
20

15 61 CCTGTATTCAGATGATAACAAAGTTGAAACTGGTGATCAAGGATATAACAGTGGTACAA 120

20 + 21 P V F A D D T K V E T G D Q G Y T V V Q  
40

40  $\Delta$

20 121 AGCAAGTATAAGAAAGCTGTTGAAACAATTACAAAAGGAATTAGATGGAAGTATAACA 180

25 41 S K Y K K A V E Q L Q K G I L D G S I T  
60

25 181 GAAATTAAAGTTCTTGAGGGAACTTTAGCATCTACTATAAAAGTAGGTTCTGAGCTT 240

30 61 E I K V F F E G T L A S T I K V G S E L  
80

30 241 AATGCAGCAGATGCAAGTAATTATTGTTACACAAGTAGATAATAAACTAGATAATTAA 300

35 81 N A A D A S K L L F T Q V D N K L D N L  
100

35 301 GGTGATGGAGATTATGTAGATTCTTAATAACTTCTCCAGGTCAAGGGGATAAAATAACT 360

40 101 G D G D Y V D F L I T S P G Q G D K I T  
120

40 361 ACAAGTAAACTTGTGCATTGAAAGATTAAACAGGTGCTTCAGCAGATGCTATAATTGCT 420

45 121 T S K L V A L K D L T G A S A D A I I A  
140

45 421 GGAACATCTTCAGCAGATGGTGTGTTACAAACTGGAGCTGCTAGTGGTTCTACTGAG 480

160 141 G T S S A D G V V T N T G A A S G S T E

481  
 ACAAAATTCAGCAGGAACAAAACCTTGCAATGTCAGCTATTTTGACACAGCATATAAGAT 540  
 5 161 T N S A G T K L A M S A I F D T A Y T D  
 180 541  
 TCATCTGAAACTGCGGTTAAGATTACTATAAAGCAGATATGAATGATACTAAATTGGT 600  
 10 181 S S E T A V K I T I K A D M N D T K F G  
 200 601  
 AAAGCAGGTGAGACAACCTTATTCAACTGGGCTTACATTGAAGATGGGTCTACAGAAAAA 660  
 15 201 K A G E T T Y S T G L T F E D G S T E K  
 220 661  
 ATTGTTAAATTAGGGGACAGTGTATTAGTATAACTAAAGCTTAAACTTACTGTT 720  
 20 221 I V K L G D S D I I D I T K A L K L T V  
 240 721  
 GTTCTGGAAAGTAAAGCAACTGTTAAGTTGCTGAAAAAACACCAAGTGCCTAGTCAA 780  
 25 241 V P G S K A T V K F A E K T P S A S V Q  
 260 781  
 CCAGTATAACAAGCTTAGAATAATAATGCTAAAGAAGAACATAGATATTGACGCT 840  
 30 261 P V I T K L R I I N A K E E T I D I D A  
 280 841  
 AGTTCTAGAAAACAGCACAGATTAGCTAAAAAATATGTTATTTAAATAACTGATTA 900  
 35 281 S S S K T A Q D L A K K Y V F N K T D L  
 300 901  
 AATACTCTTATAAAAGTATTAATGGAGATGAAGCAGATACTAATGGATAATAGAAGAA 960  
 40 301 N T L Y K V L N G D E A D T N G L I E E  
 320 961  
 GTTAGTGGAAAATATCAAGTAGTTCTTATCCAGAAGGAAAAGAGTTACAACTAAGAGT 1020  
 45 321 V S G K Y Q V V L Y P E G K R V T T K S  
 340 1021  
 GCTGCAAAGGCTCAATTGCTGATGAAAATTCACCAAGTTAACTCTTAAAGTCAGAT 1080  
 50 341 A A K A S I A D E N S P V K L T L K S D  
 360  
 ♦  
 1081  
 AAGAAGAAAGACTTAAAGATTATGTTGATGATTAAGAACATATAATGGATATTCA 1140  
 55 361 K K K D L K D Y V D D L R T Y N N G Y S  
 380

1141  
 AATGCTATAGAAGTAGCAGGAGAAAGATAGAAATAGAAACTGCAATAGCATTAAGTCAAAAA 1200  
 381 N A I E V A G E D R I E T A I A L S Q K  
 5 400  
 1201  
 TATTATAACTCTGATGATGAAATGCTATATTAGAGATTAGTTGATAATGTAGTATTG 1260  
 401 Y Y N S D D E N A I F R D S V D N V V L  
 10 420  
 1261  
 GTTGGAGGAAATGCAATAGITGATGGACTTGTAGCTTCCTTAGCTCTGAAAAGAAA 1320  
 421 V G G N A I V D G L V A S P L A S E K K  
 15 440  
 1321  
 GCTCCCTTATTATTAACCTCAAAAGATAAAATTAGATTCAAGCGTAAAGCTGAAAATAAAG 1380  
 441 A P L L L T S K D K L D S S V K A E I K  
 20 460  
 1381  
 AGAGTTATGAATATAAAGAGTACAACAGGTATAAATCTCAAGAGAAAAGCTTATTAGCT 1440  
 461 R V M N I K S T T G I N T S K K V Y L A  
 25 480  
 1441  
 GGTGGAGTTAATTCTATATCTAAAGAAGTAGAAAATGAATTAAAAGATATGGGACTTAAA 1500  
 481 G G V N S I S K E V E N E L K D M G L K  
 30 500  
 1501  
 GTTACAAGATTAGCAGGAGATGATAGATATGAAACTTCTCTAAAGATAGCTGATGAAGTA 1560  
 501 V T R L A G D D R Y E T S L K I A D E V  
 35 520  
 1561  
 GGTCTTGATAATGATAAACGCATTTGTAGTTGGAGGAACAGGATTAGCAGATGCCATAGT 1620  
 521 G L D N D K A F V V G G T G L A D A M S  
 40 540  
 1621  
 ATAGCTCCAGTTGCATCTCAATTAAGAAATGCTAATGGTAAATGGATTAGCTGATGGT 1680  
 541 I A P V A S Q L R N A N G K M D L A D G  
 45 560

1681  
GATGCTACACCAATAGTAGTTGAGATGGAAAAGCTAAACACTATAAATGATGATGTA  
5 561 D A T P I V V V D G K A K T I N D D V K K  
580  
1741  
GATTCTTAGATGATTACAAGTTGATATAATAGGTGGAGAAAACAGTGATCTAAAGAT 1800  
10 581 D F L D D S Q V D I I G G E N S V S K D  
600  
1801  
GTTGAAAATGCAATAGATGATGCTACAGGTAATCTCCAGATAGATAGTGGAGATGAT 1860  
15 601 V E N A I D D A T G K S P D R Y S G D D  
620  
1861  
AGACAAAGCAACTAATGCAAAAGTTATAAAAGAATCTCTTATTATCAAGATAACTAAAT 1920  
20 621 R Q A T N A K V I K E S S S Y Y Q D N L N  
640  
1921  
AATGATAAAAAGTAGTTAATTCTTTGAGCTAAAGATGGTTCTACTAAAGAAGATCAA 1980  
25 641 N D K K V V N F F V A K D G S T K E D Q  
660  
1981  
TTAGTTGATGCTTCTAGCAGCAGCTCAGTTGCAGCAAACCTTGGTGTAACTCTTAAATTCT 2040  
30 661 L V D A L A A A P V A A N F G V T L N S  
680  
2041  
GATGGTAAGCCAGTAGATAAAAGATGGTAAAGtATTAACGGTTCTGATAATGATAAAAAT 2100  
35 681 D G K P V D K D G K V L T G S D N D K N  
700  
2101  
AAATTAGTATCTCAGCACCTATAGTATTAGCTACTGATTCTTATCTTCAGATCaAAGT 2160  
40 701 K L V S P A P I V L A T D S L S S D Q S  
720  
2161  
GTATCTATAAGTAAGTTCTTGATAAAAGATAATGGAGAAAACCTTAGTTCAAGTTGGTAAA 2220  
45 721 V S I S K V L D K D N G E N L V Q V G K  
740  
2221 GGTATAGCTACTTCAGTTATAACAAAATGAAAGATTATTAGATATG 2268  
741 G I A T S V I N K M K D L L D M 756

## Appendix 2

SEQ ID No. 4. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 172450, PCR type 5, with translation. The putative secretory signal cleavage site ( $\Delta$ ) is indicated, and an approximation of the and site of cleavage to form the two mature SLPs ( $\blacklozenge$ ) is also indicated.

5	ATGAAAAAAAGAAATTAGCAATGGCTATGGCAGCTGTTACTGTAGTAGGTTCTGCCTGCT	60	
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+		
	1 M K K R N L A M A M A A V T V V G S A A		
10	15	CCAGTTTGCAGCAGCTTCAGATGTAATATCACTACAAGATGGTACAAATGATAAGTAT	120
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+		
	21 P V F A A A S D V I S L Q D G T N D K Y		
15	20	40	$\Delta$
20	121	ACAGTATCAAATACTAAAGCTAGTAGCTAGTAAAGGATATTTAGCAGCACAAAACCTA	180
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	41 T V S N T K A S D L V K D I L A A Q N L		
25	60	60	
25	181	ACAACAGGTGCAGTTATTTGAACAAAGATACAAAAGTTACTTCTATGATGCAAATGAG	240
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	61 T T G A V I L N K D T K V T F Y D A N E		
30	80	80	
30	241	AAAGATTCTTCAACTCCAAGTGAGATAAAAAGTTATTCTAGAACAAACTTAACTACA	300
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	81 K D S S T P T G D K K V Y S E Q T L T T		
35	100	100	
35	301	GCTAATGGAAATGAAGATTATGTAAGACAACTTAAAAAATTAGATGCGAGGAAATAT	360
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	101 A N G N E D Y V K T T L K N L D A G E Y		
40	120	120	
40	361	GCTATTATAGATTAACTTATAATAATGCTAAACTGTTGAAATTAAAGTAGTAGCTAGCAGCT	420
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	121 A I I D L T Y N N A K T V E I K V V A A		
45	140	140	
45	421	AGTGAAAAAACAGTAGTTGTATCTAGTGATGCGAAAAATGTGCAAAAGATATAGCTGAA	480
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	141 S E K T V V V S S D A K N S A K D I A E		
50	160	160	
50	481	AAATATGTGTTGAAGACAAAGACTTAGAAAAATGCACTAAAAACTATAATGCCTCAGAT	540
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	161 K Y V F E D K D L E N A L K T I N A S D		
55	180	180	

541

TTCAGTAAACTGATAGTTACTATCAAGTAGTTCTTATCCAAAAGGAAAGAGATTACAA 600

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

5       181 F S K T D S Y Y Q V V L Y P K G K R L Q

200      200 601

GGTTTCTCAACTTATAGAGCTACAAATTATAATGAAGGAACTGCATATGGTAATACACCA 660

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

10      201 G F S T Y R A T N Y N E G T A Y G N T P

220      220 ♦

15      221 V I L T L K S T S K S N L K T A V E E L

240      240 661

GTAATATTAACTCTAAATCTACTAGTAGAAGAGTAATTAAAGACTGCAGTAGAAGAGTTA 720

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

20      241 Q K L N A S Y S N T T T L A G D D R I Q

260      260 721

CAAAAATTGAATGCTAGTTATTCTAAATACTACAACTTTAGCTGGTGTACAGAAATACA 780

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

25      261 T A I E I S K E Y Y N N D G E K S D H S

280      280 841

GCTGTATGTTAAAGAGAATGTTAAAAATGTTGTATTAGTAGGTGCAAATGCACTAGTAGAT 900

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

30      281 A D V K E N V K N V V L V G A N A L V D

300      300 901

GGATTAGTTGGCGGCTCTTCTAGCAGCAGAAAAAGATGCTCCACTATTATTAACCTCAAA 960

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

35      301 G L V A A P L A A E K D A P L L T S K

320      320 961

GATAAAATTAGATTGTCAGTAAATCTGAAATAAAGAGAGTTAGACTTAAACTTC 1020

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

40      321 D K L D S S V K S E I K R V L D L K T S

340      340 1021

ACAGAAGTAACAGGAAAAACAGTTATATAGCTGGTGAGTTAATAGTGTATCTAAAGAA 1080

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

45      341 T E V T G K T V Y I A G G G V N S V S K E

360      360 1081

GTGTAAACAGAATTAGAACATGGGATTAAAAGTTGAAAGATTCTCAGGTGATGATAGA 1140

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

50      361 V V T E L E S M G L K V E R F S G D D R

380      380 1141

TATGAAACTCTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATGATAAGGCTTATGTA 1200

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

381 Y E T S L K I A G E I G L D N D K A Y V  
400  
1201  
5 GTTGGTGGAAACAGGATTAGCAGATGCCATGAGTATAGCTTCAGTTGCTTCACTAAATTA 1260  
-----+-----+-----+-----+  
401 V G G T G L A D A M S I A S V A S T K L  
420  
1261  
10 GATGGTAATGGTGTGAGATAAGAACAAATGGACATGCTACTCCAATAGTTGTTGAGAT 1320  
-----+-----+-----+-----+  
421 D G N G V V D R T N G H A T P I V V V D  
440  
1321  
15 GAAAAGCTGATAAAAATATCTGATGACTTAGATAGTTCTTAGGAAGCGCTGATGAGAT 1380  
-----+-----+-----+-----+  
441 G K A D K I S D D L D S F L G S A D V D  
460  
1381 ATAATAGTGGATTGCAAGTGTATCTGAAAAGATGGAAGAGCTATATCAGATGCTACT  
1440  
20 461 I I G G F A S V S E K M E E A I S D A T  
480  
1441  
25 GGTAAAGGCCTTACAAGAGTTAAAGGCACGATAGACAAGACACTAATCTGAAGTTATA 1500  
-----+-----+-----+-----+  
481 G K G V T R V K G D D R Q D T N S E V I  
500  
1501  
30 AAAACATATTATGCTAATGATACTGAAATAGCTAAAGCTGCAGTTAGATAAAAGATTCA 1560  
-----+-----+-----+-----+  
501 K T Y Y A N D T E I A K A A V L D K D S  
520  
1561  
35 GGTGCTCAAGTAGTGATGCAGGAGTATTAATTCTATGCTAAAGATGGATCTACA 1620  
-----+-----+-----+-----+  
521 G A S S S D A G V F N F Y V A K D G S T  
540  
1621  
40 AAAGAAGATCAATTAGTTGATGCATTAGCAGTAGGAGCTGTTGCTGGATATAAAGCTTGCT 1680  
-----+-----+-----+-----+  
541 K E D Q L V D A L A V G A V A G Y K L A  
560

1681  
CCAGTTGTATTAGCTACTGATTCTTATCTCTGATCAATCGGTTGCTATAAGCAAAGTT 1740  
5 561 P V V L A T D S L S S D Q S V A I S K V  
580  
1741  
GTAGGAGAAAAATATTCTAAAGATTTAACACAAGTTGGTCAAGGAATAGCTAATTCAAGTT 1800  
10 581 V G E K Y S K D L T Q V G Q G I A N S V  
600  
1801 ATAAACAAAATGAAAGATTTATTAGATATG 1830  
15 601 I N K M K D L L D M 610

## Appendix 3

SEQ ID No. 5. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 170324, PCR type 12, with translation. The putative secretory signal cleavage site (A) and site of cleavage to form the two mature SLPs (♦) are indicated.

5	1 ATGAAATAAGAAAAATAGCAATAGCTATGTCAGGTTAACAGTTAGCTCGGCTGCT 10           60 -----+-----+-----+-----+-----+----- 1 M N K K N I A I A M S G L T V L A S A A
15	20           61 CCTTTTGTGCTGCAACTACTGGAACACAAGGTTACTGTAGTTAAAACGACTGGAAA 120 -----+-----+-----+-----+-----+----- 21 P V F A A T T G T Q G Y T V V K N D W K
20	40           41 AAAGCAGTAAACAAATTACAAGATGGACTAAAAGATAATAGTATAGGAAAGATAACTGTA 180 -----+-----+-----+-----+-----+----- 41 K A V K Q L Q D G L K D N S I G K I T V
25	60           61 181           TCTTTTAATGATGGGGTTGTGGGTGAAGTAGCTCTAAAGTGCTAATAAGAAAGCGGAC 240 -----+-----+-----+-----+-----+----- 61 S F N D G V V G E V A P K S A N K K A D
30	80           81 241           AGAGATGCTGCAGCTGAGTTATATAATCTTGTAAACACTCAATTAGATAATTAGGT 300 -----+-----+-----+-----+-----+----- 81 R D A A A E K L Y N L V N T Q L D K L G
35	100           101 301           GATGGAGATTATGTTGATTTCTGTAGATTATAATTAGAAAACAAAATAACTAAT 360 -----+-----+-----+-----+-----+----- 101 D G D Y V D F S V D Y N L E N K I I T N
40	120           121 361           CAAGCAGATGCAGAAGCAATTGTTACAAAGTTAAATTCACTTAATGAGAAAACCTTATT 420 -----+-----+-----+-----+-----+----- 121 Q A D A E A I V T K L N S L N E K T L I
45	140           141 421           GATATAGCAACTAAAGATACTTTGGAATGGTTAGTAAACACAAGATAGTGAAGGTAAA 480 -----+-----+-----+-----+-----+----- 141 D I A T K D T F G M V S K T Q D S E G K
	160

481  
 ATGTTGCTGCAACAAAGGCACCTAAAGTTAAAGATGTTGCTACATTGGTTGAAGTCT 540

5 161 N V A A T K A L K V K D V A T F G L K S

180 541 GGTGGAAGCGAAGATACTGGATATGTTGTTGAAATGAAAGCAGGAGCTGTAGAGGATAAG 600

10 181 G G S E D T G Y V V E M K A G A V E D K

200 601 TATGGTAAAGTGGAGATAGTACGGCAGGTATTGCAATAAATCTCTAGTACTGGACTT 660

15 201 Y G K V G D S T A G I A I N L P S T G L

220 661 GAATATGCAGGTAAGGAACAACAAATTGATTTAATAAAACTTAAAGTTGATGTAAACA 720

20 221 E Y A G K G T T I D F N K T L K V D V T

240 721 GGTGGTCAACACCTAGTGTAGCTGAACTAAAGATGATACTGAT 780

25 241 G G S T P S A V A V S G F V T K D D T D

260 781 TTAGCAAAATCAGGTACTATAATGTAAGAGTTATAATGCAAAAAGAAGAATCAATTGAT 840

30 261 L A K S G T I N V R V I N A K E E S I D

280 841 ATAGATGCAAGCTCATATACATCAGCTGAAAAATTAGCTAAAAGATATGTATTCATCCA 900

35 281 I D A S S Y T S A E N L A K R Y V F D P

300 901 GATGAAATTCTGAAGCATATAAGGCAATAGTAGCATTACAAAATGATGGTATAGAGTCT 960

40 301 D E I S E A Y K A I V A L Q N D G I E S

320 961 RACTTAGTTCAAGTTAGTTATGGAAAAATATCAAGTGATTTTATCCAGAAGCTAAAGAGA 1020

45 321 N L V Q L V N G K Y Q V I F Y P E G K R

340 1021 TTAGAAACTAAATCAGCAAATGATAATAGCTAGTCAGGATACACCAGCTAAAGTAGTT 1080

50 341 L E T K S A N D T I A S Q D T P A K V V

360 ♦

1081  
 ATAAAAAGCTAATAAATTAAAAGATTAAAAGATTATGAGATGATTAAAAACATATAAT 1140

361 I K A N K L K D L K D Y V D D L K T Y N  
 380  
 1141  
 5 AATACTTATTCAAATGTTAACAGTAGCAGGAGAAGATAGAAACTGCTATAGAA 1200  
 381 N T Y S N V V T V A G E D R I E T A I E  
 400  
 1201  
 10 TTAAGTAGTAAATATTATAATTCTGATGATAAAATGCAATACTGATAAAGCAGTTAAT 1260  
 401 L S S K Y Y N S D D K N A I T D K A V N  
 420  
 1261  
 15 GATATAGTATTAGTGGATCTACATCTATAGTTGATGGTCTTGGCATCACCATTAGCT 1320  
 421 D I V L V G S T S I V D G L V A S P L A  
 440  
 1321  
 20 TCAGAAAAACAGCTCCATTATTATAACTCAGAAAGATAATTAGATTATCAGTAAAGA 1380  
 441 S E K T A P L L L T S K D K L D S S V K  
 460  
 1381  
 25 TCTGAAATAAGAGAGTTATGAACTTAAAGAGTGACACTGGTATAAAACTTCTAAAAAA 1440  
 461 S E I K R V M N L K S D T G I N T S K K  
 480  
 1441  
 30 GTTATTAGCTGGTGGAGTTAATTCTATCTAAAGATGTAGAAAATGAATTGAAAAAC 1500  
 481 V Y L A G G V N S I S K D V E N E L K N  
 500  
 1501  
 35 ATGGGTCTTAAAGTTACTAGATTATCAGGAGAAGACAGATACGAAACTCTTAGCAATA 1560  
 501 M G L K V T R L S G E D R Y E T S L A I  
 520  
 1561  
 40 GCTGATGAAATAGGTCTTGATAATGATAAAGCATTGTAGTTGGTAGCTGGATTAGCA 1620  
 521 A D E I G L D N D K A F V V G G T G L A  
 540  
 1621  
 45 GATGCTATGAGTATAGCTCCAGTTGCTTCTCAACTAAAGATGGAGATGCTACTCCAATA 1680  
 541 D A M S I A P V A S Q L K D G D A T P I  
 560  
 50  
 1681  
 561 V V V D G K A K E I S D D A K S F L G T  
 580  
 1741  
 55 TCTGATGTTGATATAATAGGTGGAAAAATAGCGTATCTAAAGAGATTGAAGAGTCATAA 1800

581 S D V D I I G G K N S V S K E I E E S I  
600  
1801 GATAGTCAACTGGAAAAACTCCAGATAGAATAAGTGGAGATGATAGACAAGCAACTAAT 1860  
5 601 D S A T G K T P D R I S G D D R Q A T N  
620  
1861 GCTGAAGTTAAAGAAGATGATTACAGATGGTGAAGTTGTGAATTACTTGTT 1920  
10 621 A E V L K E D D Y F T D G E V V N Y F V  
640  
1921 GCAAAAGATGGTCTACTAAAGAAGATCAATTAGTAGATGCCTAGCAGCACCAATA 1980  
15 641 A K D G S T K E D Q L V D A L A A A P I  
660  
1981 GCAGGTAGATTTAAGGAGTCTCCAGCTCCAATCATACTAGTACTGATACTTTATCTTCT 2040  
20 661 A G R F K E S P A P I I L A T D T L S S  
680  
2041 GACCCAAATGTAGCTGTAAGTAAAGCAGTTCTAAAGATGGTGGACTAACTTAGTTCAA 2100  
25 681 D Q N V A V S K A V P K D G G T N L V Q  
700  
2101 GTAGGTAAAGGTATAGCTTCTTCAGTTATAAACAAAATGAAAGATTTATTAGATATG  
2157  
30 701 V G K G I A S S V I N K M K D L L D M  
719

## Appendix 4

SEQ ID No 6. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 171448, PCR type 12, with translation. The putative secretory signal cleavage site (A) and site of cleavage to form the two mature SLPs (♦) are indicated.

1 ATGAAATAAGAAAAATATAGCAATAGCTATGTCAGGTTAACAGTTTAGCTTCGGCTGCT 60  
 10 +-----+-----+-----+-----+-----+-----+  
 1 M N K K N I A I A M S G L T V L A S A A  
 20 +-----+-----+-----+-----+-----+-----+  
 61 CCTGTTTGTGCTGCAACTACTGGAACACAAGGTTAACTGTAGTTAAAACGACTGGAAA 120  
 15 +-----+-----+-----+-----+-----+-----+  
 21 P V F A A T T G T Q G Y T V V K N D W K  
 40 +-----+-----+-----+-----+-----+-----+  
 121 Δ  
 20 AAAGCAGTAAACAAATTACAAGATGGACTAAAGATAATAGTATAGGAAAGATAACTGTA 180  
 25 +-----+-----+-----+-----+-----+-----+  
 41 K A V K Q L Q D G L K D N S I G K I T V  
 60 +-----+-----+-----+-----+-----+-----+  
 181 TCTTTAATGATGGGGTTGGGTGAAGTAGCTCTAAAGTGTAAATAAGAAAGCGGAC 240  
 25 +-----+-----+-----+-----+-----+-----+  
 61 S F N D G V V G E V A P K S A N K K A D  
 80 +-----+-----+-----+-----+-----+-----+  
 241 AGAGATGCTGCAGCTGAGAAGTTATATAATCTTGTAAACACTCAATTAGATAATTAGGT 300  
 30 +-----+-----+-----+-----+-----+-----+  
 81 R D A A A E K L Y N L V N T Q L D K L G  
 100 +-----+-----+-----+-----+-----+-----+  
 301 GATGGAGATTATGTTGATTTCTGTAGATTATAATTAGAAAACAAAATAAACTAAT 360  
 35 +-----+-----+-----+-----+-----+-----+  
 101 D G D Y V D F S V D Y N L E N K I I T N  
 120 +-----+-----+-----+-----+-----+-----+  
 361 CAAGCAGATGCGAGCAAGCAATTGTTACAAAGTTAAATTCACTTAATGAGAAAACCTTATT 420  
 40 +-----+-----+-----+-----+-----+-----+  
 121 Q A D A E A I V T K L N S L N E K T L I  
 140 +-----+-----+-----+-----+-----+-----+  
 421 GATATGCAACTAAAGATACTTTGGATGGTTAGTAAACACAAGATACTGGAGGTAAA 480  
 45 +-----+-----+-----+-----+-----+-----+  
 160 141 D I A T K D T F G M V S K T Q D S G G K

481  
AATGTTGCTGCAACAAAGGCACCTTAAGTTAAAGATGTTGCTACATTGGTTGAAGTCT 540

5 161 N V A A T K A L K V K D V A T F G L K S  
180 541 GGTGGAAGCGAAGATACTGGATATGTTGTTGAAATGAAAGCAGGAGCTGTAGAGGATAAG 600

10 181 G G S E D T G Y V V E M K A G A V E D K  
200 601 TATGGTAAAGTTGAGATAGTACGGCAGGTATTGCAATAAATCTCTAGTACTGGACTT 660

15 201 Y G K V G D S T A G I A I N L P S T G L  
220 661 GAATATGCGAGGTAAGGAAACAACATTGATTTAAATAAAACTTAAAGTTGATGTAACA 720

20 221 E Y A G K G T T I D F N K T L K V D V T  
240 721 GGTGGTTCAACACCTAGTGTAGCTGTAAAGTGGTTTGTAACTAAAGATGATACTGAT 780

25 241 G G S T P S A V A V S G F V T K D D T D  
260 781 TTAGCAAATCAGGTACTATAATGTAAGAGTTAAATGCAAAAGAAGAATCAATTGAT 840

30 261 L A K S G T I N V R V I N A K E E S I D  
280 841 ATAGATGCAAGCTCATATACATCAGCTGAAAATTAGCTAAAGATATGTATTTGATCCA 900

35 281 I D A S S Y T S A E N L A K R Y V F D P  
300 901 GATGAAATTCTGAAGCATATAAGGCAATAGTAGCATTACAAAATGATGGTATAGAGTCT 960

40 301 D E I S E A Y K A I V A L Q N D G I E S  
320 961 ATTTCAGTTCAAGTTATGGAAAAATATCAAGTGTATTTTATCCAGAAGGTAAAAGA 1020

45 321 N L V Q L V N G K Y Q V I F Y P E G K R  
340 1021 TTAGAAAACCTAAATCAGCAATGATACAATAGCTAGTCAGATAACCCAGCTAAAGTAGTT 1080

50 341 L E T K S A N D T I A S Q D T P A K V V  
360

◆

1081  
55 ATAAAAGCTAATAAAATTAAAGATTTAAAGATTATGAGATGATTAAAAACATATAAT 1140

361 I K A N K L K D L K D Y V D D L K T Y N  
 380  
 1141  
 5 AATACTTATTCAAATGTTAACAGTAGCAGGAGAAGATAGAAACTGCTATAGAA 1200  
 381 N T Y S N V V T V A G E D R I E T A I E  
 400  
 1201  
 10 TTAAGTAGTAAATTATAATTCTGATGATAAAATGCAATAACTGATAAAGCAGTTAAT 1260  
 401 L S S K Y Y N S D D K N A I T D K A V N  
 420  
 1261  
 15 GATATAGTATTAGTTGGATCTACATCTATAGTTGATGGCTTGTCATCACCATTAGCT 1320  
 421 D I V L V G S T S I V D G L V A S P L A  
 440  
 1321  
 20 TCAGAAAAACAGCTCCATTATTATTAGCTCAAAAGATAATTAGATTCATCAGTAAAA 1380  
 441 S E K T A P L L L A S K D K L D S S S V K  
 460  
 1381  
 25 TCTGAAATAAAAGAGAGTTATGAACTTAAAGAGTGACACTGGTATAAAACTTCTAAAAA 1440  
 461 S E I K R V M N L K S D T G I N T S K K  
 480  
 1441  
 30 GTTATTAGCTGGGAGTTAATTCTATCTAAAGATGTAGAAAATGAATTGAAAAAC 1500  
 481 V Y L A G G V N S I S K D V E N E L K N  
 500  
 1501  
 35 ATGGGTCTTAAAGTTACTAGATTATCAGGAGAAGACAGATACGAAACTCTTAGCAATA 1560  
 501 M G L K V T R L S G E D R Y E T S L A I  
 520  
 1561  
 40 GCTGATGAAATAGGTCTTGATAATGATAAAGCATTGTAGTTGGTGGTACTGGATTAGCA 1620  
 521 A D E I G L D N D K A F V V G G T G L A  
 540  
 1621  
 45 GATGCTATGAGTATAGCTCCAGTTGCTTCTCAACTTAAAGATGGAGATGCTACTCCAATA 1680  
 541 D A M S I A P V A S Q L K D G D A T P I  
 560  
 50  
 1681  
 580  
 55 GTAGTTGAGATGAAAAGCAAAGAAATAAGTGTGATGCTAAGAGTTCTTAGGAAC 1740  
 561 V V V D G K A K E I S D D A K S F L G T  
 580  
 1741  
 TCTGATGTTGATATAATAGGTGGAAAAAATAGCGTATCTAAAGAGATTGAACAGTCATAA 1800

581 S D V D I I G G K N S V S K E I E E S I  
600  
1801  
5 GATAGTGCACACTGGAAAAACTCCAGATAGATAAGTGGAGATGATAGACAAGCAACTAAT 1860  
-----+-----+-----+-----+-----+  
601 D S A T G K T P D R I S G D D R Q A T N  
620  
1861  
10 GCTGAAGTTAAAGAAGATGATTACAGATGGTGAAGTTGTGAATTACTTTGTT 1920  
-----+-----+-----+-----+-----+  
621 A E V L K E D D Y F T D G E V V N Y F V  
640  
1921  
15 GCAAAAGATGGTTACTAAAGAACATCAATTAGTAGATGCCTTAGCAGCACCAATA 1980  
-----+-----+-----+-----+-----+  
641 A K D G S T K E D Q L V D A L A A A P I  
660  
1981  
20 GCAGGTAGATTTAAGGAGTCTCCAGCTCCAATCATACTAGCTACTGATACTTATCTCT 2040  
-----+-----+-----+-----+-----+  
661 A G R F K E S P A P I I L A T D T L S S  
680  
2041  
25 GACCCAAAATGTAGCTGTAAGTAAAGCAGTTCTAAAGATGGTGGAACTAACTTAGTC 2100  
-----+-----+-----+-----+-----+  
681 D Q N V A V S K A V P K D G G T N L V Q  
700  
2101 GTAGGTAAAGGTATAGCTTCTCAGTTATAAACAAAATGAAAGATTATTAGATATG  
2157  
30  
701 V G K G I A S S V I N K M K D L L D M  
719

## Appendix 5

SEQ ID No. 7. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 171862, PCR type 17, with translation. The putative secretory signal cleavage site ( $\Delta$ ) and site of cleavage to form the two mature SLPs ( $\diamond$ ) are indicated.

```

1 ATGAATAAGAAAACCTAGCAATGGCTATGGCAGCAGTTACTGTTGGGCTCTGCAGCG 60
10 +-----+-----+-----+-----+
1 M N K K N L A M A M A A V T V V G S A A
20 +-----+
61 CCAATATTGCGAGATAGTACTACCCCAGGTTATACTGTAGTGAAAAATGATTGGAAAAA 120
15 +-----+-----+-----+-----+
21 P I F A D S T T P G Y T V V V K N D W K K
40 +-----+
20 121 GCACTAAACAACTTACAAGATGGGTTGAAAAAATAAAACTATATCAAACAAATAAGGTGCT 180
41 A V K Q L Q D G L K N K T I S T I K V S
60 +-----+
25 181 TTTAATGGAAACTCTGTTGGAGAAGTTACACCAGCCAGTCTGGAGCAAAAAAAGCAGAT 240
61 F N G N S V G E V T P A S S S G A K K A D
80 +-----+
30 241 AGAGATGCTGCAGCTGAAAAGTTATATAATTAGTAATACACAATTAGATAAACTAGGT 300
81 R D A A A E K L Y N L V N T Q L D K L G
100 +-----+
35 301 GATGGAGATTACGTTGACTTGAAGTAACCTATAATTAGCTACTCAAATAATTACAAAA 360
101 D G D Y V D F E V T Y N L A T Q I I T K
120 +-----+
40 361 GCAGAAAGCAGAGGCCAGTTCTTACAAAATTACAACAAATAATGATAAAAGTACTTATAAT 420
121 A E A E A V L T K L Q Q Y N D K V L I N
140 +-----+
45 421 TCTGAAACRGATACAGTAAAGGTATGGTATCTGATACACAAGTTGATAGCAAAATGTT 480
141 S A T D T V K G M V S D T Q V D S K N V
160 +-----+

```

481  
GCAGCTAACCCACTTAAAGTTAGTGATATGTACAAATACCATCTGCTATTACTGGAAAGT 540  
-----+-----+-----+-----+-----+-----+  
5 161 A A N P L K V S D M Y T I P S A I T G S  
180 541  
GATGATTCTGGGTATAGTATTGCTAAACCAACAGAAAAGACTACAAgTTTATTGTATGGT 600  
-----+-----+-----+-----+-----+-----+  
10 181 D D S G Y S I A K P T E K T T S L L Y G  
200 601  
ACGGTTGGTGTATGCCAAGTCAGGTAAGCAATAACAGTAGATACAGCTTCAAATGAAGCT 660  
-----+-----+-----+-----+-----+-----+  
15 201 T V G D A T A G K A I T V D T A S N E A  
220 661  
TTGCTGGAAATGAAAGGTTATTGACTACAAATAATCATTCAAACCAACTGTACAAGGA 720  
-----+-----+-----+-----+-----+-----+  
20 221 F A G N G K V I D Y N K S F K A T V Q G  
240  
  
721  
GATGGAACAGTTAACAGAACAGCGGGGTTGACTTAAAGATGCAAGTGTATGGCTGCAACA 780  
-----+-----+-----+-----+-----+-----+  
25 241 D G T V K T S G V V L K D A S D M A A T  
260 781  
GGTACTATAAAAGTTAGAGTTACAAAGTGCAGGAAAGAAGAATCTATTGATGTGGATTCAAGT 840  
-----+-----+-----+-----+-----+-----+  
30 261 G T I K V R V T S A K E E S I D V D S S  
280 841  
TCATATATTAGTGCTGAAATTTAGCTAAAAAATATGTATTTAACCTAAAGAGGTTCT 900  
-----+-----+-----+-----+-----+-----+  
35 281 S Y I S A E N L A K K Y V F N P K E V S  
300 901  
GAAGCTTATAATGCCAATAGTTGCATTACAAAATGATGGAATAGAAATCTGATTTAGTACAA 960  
-----+-----+-----+-----+-----+-----+  
40 301 E A Y N A I V A L Q N D G I E S D L V Q  
320 961  
TTAGTTAATGGAAAATATCAAGTTATTTCTATCCAGAAGGAAAAGATTAGAAACTAAA 1020  
-----+-----+-----+-----+-----+-----+  
45 321 L V N G K Y Q V I F Y P E G K R L E T K  
340

1021 TCTGCAGATATAATAGCTGATGCGAGATAGTCAGCTAAATAACTATAAAGCTAATAAA  
 1080  
 341 S A D I I A D A D S P A K I T I K A N K  
 360  
 ♦  
 1081  
 TTAAGGATTTAAAGATTATGTAGATGATTTAAAACATAACAATAACTTACTCAAAT 1140  
 361 L K D L K D Y V D D L K T Y N N T Y S N  
 380  
 1141  
 GTTGTAACTAGCAGGAGAAGATAAGATAAACTGCTATAGAATTAAGTAGTAAATAT 1200  
 381 V V T V A G E D R I E T A I E L S S K Y  
 400  
 1201  
 TATAATTCTGATGATAAAAATGCAATAACTGATGATGCACTTAATAATAGTATTAGTT 1260  
 401 Y N S D D K N A I T D D A V N N I V L V  
 420  
 1261  
 GGATCTACATCTATAGTTGATGGCTTGTGCACTACCATTAGCTCAGAAAAAACAGCT 1320  
 421 G S T S I V D G L V A S P L A S E K T A  
 440  
 1321  
 CCATTATTAACTCAAAAGATAATTAGATTCATCAGTAAATCTGAGATAAAAAGA 1380  
 441 P L L L T S K D K L D S S V K S E I K R  
 460  
 1381  
 GTTATGAACCTAAAGAGTGATACTGGTATAAAACTCTAAAAAGTTTATTTAGCTGGT 1440  
 461 V M N L K S D T G I N T S K K V Y L A G  
 480  
 1441  
 GGAGTTAATTCTATCTAAAGATGTAGAGATGAATTGAAAAATATGGCCTTAAAGTT 1500  
 481 G V N S I S K D V E D E L K N M G L K V  
 500  
 1501  
 ACTAGATTATCAGGAGAACAGATACGAAACTCTTTAGCAATAGCTGATGAAATAGGT 1560  
 501 T R L S G E D R Y E T S L A I A D E I G  
 520  
 1561  
 CTTGATAATGATAAAAGCATTGATGGTGGTACTGGATTGGCAGATGCTATGAGTATA 1620  
 521 L D N D K A F V V G G T G L A D A M S I  
 540  
 1621  
 GCTCCAGTTGCTCTCAACTAAAGATGGAGATGCTACTCCAATAGTTGAGATGGA 1680

541 A P V A S Q L K D G D A T P I V V V D G  
560  
1681 AAAGCAAAAGAAAATAAGTGTATGCTAAGAGTTCTTAGGAACCTCTGATGTTGATATA 1740  
5 K A K E I S D D A K S F L G T S D V D I  
580  
1741 ATAGGTGGAAAAAAATAGCGTATCTAAAGAGATTGAAGAGTCATAGATAGTGCACACTGGA 1800  
10 I G G K N S V S K E I E E S I D S A T G  
600  
1801 AAAACTCCAGATAAGATAAGTGGAGATGACAGACAAGCAACTAATGCTGAAGTTTTAAAA 1860  
15 K T P D R I S G D D R Q A T N A E V L K  
620  
1861 GAAGATGATTATTCAAAGATGGTGAAGTTGTGAATTACTTGTGCAAAGATGGTCT 1920  
20 E D D Y F K D G E V V N Y F V A K D G S  
640  
1921 ACTAAAGAGATCAATTAGTAGATGCTTAGCAGCAGCACCAATAGCAGGTAGATTTAAG 1980  
25 T K E D Q L V D A L A A A P I A G R F K  
660  
1981 GAGTCTCCAGCTCCAATCATACTAGCTACTGATACTTTATCTCTGACCAAAATGTAGCT 2040  
30 E S P A P I I L A T D T L S S D Q N V A  
680  
2041 GTAAGTAAAGCAGTTCTAAAGATGGTGAACTAACCTAGTTCAAGTAGGTAAAGGTATA 2100  
35 V S K A V P K D G G T N L V Q V G K G I  
700  
2101 GCTTCTTCAGTTATAAACAAAATGAAAGATTTATTAGATATGTAA 2145  
40 A S S V I N K M K D L L D M \* 715

## Appendix 6

SEQ ID No 8. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 173644, PCR type 31, with translation. The putative secretory signal cleavage site ( $\Delta$ ) and site of cleavage to form the two mature SLPs ( $\blacklozenge$ ) are indicated.

	1	
5	ATGAATAAGAAGGATATAGCAATAGCTATGTCAGGATTAAACAGTATTAGCTTCTGCAGCA	60
10	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	1 M N K K D I A I A M S G L T V L A S A A	
15	20	
	61	
	CCTGTATTCGCTGCTAGTAGTTACAGCAGATTATAATTACTGTAGTGCAAGGGAAAA	120
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	21 P V F A A S S F T A D Y N Y T V V Q G K	
20	40	
	121	
	TATCAAAAAGTTATAACTGGATTACAAGATGGTTAAAAATGGAAAAATAACAAAATATT	180
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	41 Y Q K V I T G L Q D G L K N G K I T N I	
25	60	
	181	
	GATGTAATTTGATGGAAGTTCAATTGGTGAGGTAGTGCCAGGTTCTGATGCTGCAGCT	240
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	61 D V I F D G S S I G E V V P G S D A A A	
30	80	
	241	
	GCAGCTACTAATTTAAAAGTTAGTTAGTGATGATAAGTTAGATAACTTAGGTGATGGAAAA	300
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	81 A A T K L K S L V D D K L D N L G D G K	
35	100	
	301	
	TACGTTCAATTAAATGTTACTTATACTACTAAATCTATAATAACTAAAGCAGAATTAAAA	360
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	101 Y V Q F N V T Y T T K S I I T K A E L K	
40	120	
	361	
	AATTATTATAATCAATTAGAAAGTAGTAAAGATAGAATACTTATAGGAAATGAACCTCAA	420
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	121 N Y Y N Q L E S S S K D R I L I G N E P Q	
45	140	
	421	
	GATACAGGAACTAAAGGTCTTATAAAAGCTGATACTGATGGTACTACTGCTGTTGCAGCA	480
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
	141 D T G T K G L I K A D T D G T T A V A A	
50	160	
	481	
	GCTGCACCATGAAATTATCAGATATTTACGTTAGTTATGATGAAGTAACAGGTGA	540
	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
55	180	
	161 A A P L K L S D I F T F S Y D E V T G V	

541  
 CTTAAAGCAGAACCAACAAGTAAAGTAAGCGCTGGTAAAGTCAAGGTCTAAAAATATGGA 600  
 181 L K A E P T S K V S A G K V Q G L K Y G  
 5 200  
 601  
 AATAACAGGAGCAACTAACTATACCTCTGGAGCTGAAATATCTGTTCTACTACAGGCTTA 660  
 201 N T G A T N Y T S G A E I S V P T T G L  
 10 220  
 661  
 ACATTAACGTGATACAACGTGAAACACAGATGTAATATTCTGATGTTATGAGTGCA 720  
 221 T L T A D T T A T T D V N I S D V M S A  
 15 240  
 721  
 TTAAATTTAATGGTACTGATACGATTAGTGGATTCCCAGCTGGTTCATCAGCTTCACT 780  
 241 F K F N G T D T I S G F P A G S S A S T  
 20 260  
 781  
 CTAGAGCAAGTATAAAAGTAATAATGCAAAGAAGAATCTATAGATGTTGATTCAAGT 840  
 261 L R A S I K V I N A K E E S I D V D S S  
 25 280  
 841  
 TCACATAGAACAGCTGAAGATTAGCTGAAAAATATGTATTAAACCAAGAGATGTGAAT 900  
 281 S H R T A E D L A E K Y V F K P E D V N  
 30 300  
 901  
 AAAACTTATGAGGCCACTGACTGATTTATATAAGAAGGTATAACAGTAATCTTACT 960  
 301 K T Y E A L T D L Y K E G I T S N L I T  
 35 320  
 961  
 CAAGATGGTGGAAAATATCAAGTTGTTTATTCCTCAAGGAAGAGATTAACATCAA 1020  
 321 Q D G G K Y Q V V L F A Q G K R L T T K  
 40 340  
 1021  
 GGAGCAACTGGAACTTTAGCAGATGAAAATCTCCCTCTTAAAGTAACAATAAGCAGAT 1080  
 341 G A T G T L A D E N S P L K V T I K A D  
 45 360  
 ♦  
 1081  
 AAAGTAAAAGACTTAAAGATTATGTTGAAGATTTAAAAAATGCTAACATGGATATTCA 1140  
 361 K V K D L K D Y V E D L K N A N N G Y S  
 380  
 1141  
 AATTCTGTTGTTGAGCTGAGATAGAAATAGAAACAGCAATAGAGTTAAGTAGCAAA 1200  
 55

381 N S V V V A G E D R I E T A I E L S S K  
 400  
 1201 TACATATACTCTGATGATGACAATGCAATAACTAAAGATCCAGTTAACATGTTGTTTA 1260  
 5  
 401 Y Y N S D D D N A I T K D P V N N V V L  
 420  
 1261 GTTGGTTCTCAAGCTGTAGTTGATGGGCTTGTAGCTCACCTTAGCATCTGAAAAAGA 1320  
 10  
 421 V G S Q A V V D G L V A S P L A S E K R  
 440  
 1321 GCTCCTTTACTATTAACTTCAGCGGGAAAATTAGATTCAAGTGTTAAAGCTGAGTTGAAA 1380  
 15  
 441 A P L L L T S A G K L D S S V K A E L K  
 460  
 1381 AGAGTAATGGATTAAATCTACACACAGGTGTAACACTTCTAAAAAGTTACTTAGCT 1440  
 20  
 461 R V M D L K S T T G V N T S K K V Y L A  
 480  
 1441 GTGGAGTAACCTATATCTAAAGATGTAGAAAATGAATTAAAAGATATGGGACTTAAA 1500  
 25  
 481 G G V N S I S K D V E N E L K D M G L K  
 500  
 1501 GTTACAAGATTATCAGGAGATGATAGATGAAACTCTTTAGCTATAGCTGATGAAATA 1560  
 30  
 501 V T R L S G D D R Y E T S L A I A D E I  
 520  
 1561 GGTCTTGATAATGATAAGCTTTGTAGTTGGAGGAACAGGATTAGCGGATGCTATGAGT 1620  
 35  
 521 G L D N D K A F V V G G T G L A D A M S  
 540  
 1621 ATAGCTCCAGTTGCTCTCAATTAAAGAAAACCTCAAATGGAGAACTTGACTTAAAGGTGAT 1680  
 40  
 541 I A P V A S Q L R N S N G E L D L K G D  
 560  
 45  
 1681 GCAACTCCAATAGTAGTTGTTGATGGAAAAGCTAAAGATATAATTCTGAAGTAAAGAT 1740  
 580  
 561 A T P I V V V D G K A K D I N S E V K D  
 50  
 1741 TTCTTAGATGATTCAACAGTTGATATAATAGGTGGTAAATAGTGTCTAAAGAAGTA 1800  
 581 F L D D S Q V D I I G G V N S V S K E V  
 600  
 1801 ATGGAAGCAATAGATGATGCTACTGGAAAATCACCTGAGAGATATACTGGAGAAGATAGA 1860  
 55

601 M E A I D D A T G K S P E R Y S G E D R  
620  
1861 CAAGCAACAAATGCTAAAGTTATAAAAGAAGATGATTCTTAAATGGAGAAGTTACA 1920  
5 621 Q A T N A K V I K E D D F F K N G E V T  
640  
1921 AACTCTTGAGCTAAAGATGGTCAACTAAAGAAGATCAATTAGTAGATGCTTAGCA 1980  
10 641 N F F V A K D G S T K E D Q L V D A L A  
660  
1981 GGTGCTGCAATTGCTGGTAACTTGGTGAACAGTAGATAATGAAGGAAAACCTACAGTT 2040  
15 661 G A A I A G N F G V T V D N E G K P T V  
680  
2041 GCTGATAAAAAGCTCTCCAGCACCAATTGTTAGCAACAGATCTTATCTCTGAT 2100  
20 681 A D K K A S P A P I V L A T D S L S S D  
700  
2101 CAAAATGTAGCTATAAGTAAAGCTGTAATGATGACGCTAATACTAAGAATCTAGTCAA 2160  
25 701 Q N V A I S K A V N D D A N T K N L V Q  
720  
2161 GTGGTAAAGGTATAGCTACTTCAGTTGTAAGTAAAATAAAAGATTATTAGATATG  
2217  
30 721 V G K G I A T S V V S K I K D L L D M  
739

## Appendix 7

SEQ ID No 9. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 170444, PCR type 46, with translation. The putative secretory signal cleavage site ( $\Delta$ ) and site of cleavage to form the two mature SLPs ( $\blacklozenge$ ) are indicated.

1 ATGAAATAAGAAAAATATAGCAATAGCTATGTCAGGTTAACAGTTAGCTTCGGCTGCT 60  
 10 -----+-----+-----+-----+-----+-----+  
 1 M N K K N I A I A M S G L T V L A S A A  
 20 61  
 CCTGTTTGCTGCAACTACTGGAACACAAAGGTTACTGTAGTTAAAACGACTGGAAA 120  
 15 -----+-----+-----+-----+-----+  
 21 P V F A A T T G T Q G Y T V V K N D W K  
 40 Δ  
 20 121  
 AAAGCAGTAAACAAATTACAAGATGGACTAAAAGATAATAGTATAGGAAAGATAACTGTA 180  
 25 -----+-----+-----+-----+  
 41 K A V K Q L Q D G L K D N S I G K I T V  
 60 181  
 TCTTTAATGATGGGTTGTTGGTGAAGTAGCTCCTAAAGTGCTAATAAGAAAGCGGGAC 240  
 25 -----+-----+-----+-----+  
 61 S F N D G V V G E V A P K S A N K K A D  
 80 241  
 AGAGATGTCAGCTGAGAAGTTATATAATCTTGTAAACACTCAATTAGATAATTAGGT 300  
 30 -----+-----+-----+-----+  
 81 R D A A A E K L Y N L V N T Q L D K L G  
 100 301  
 GATGGAGATTATGTTGATTTCTGTAGATTATAATTAGAAAAAAAATAACTAAT 360  
 35 -----+-----+-----+-----+  
 101 D G D Y V D F S V D Y N L E K K I I T N  
 120 361  
 CAAGCAGATGCGAGCAAGCAATTGTTACAAGTTAAATTCACTTAATGAGAAAACCTTATT 420  
 40 -----+-----+-----+-----+  
 121 Q A D A E A I V T K L N S L N E K T L I  
 140 421  
 GATATAGCAACTAAAGATACTTTGGAATGGTAGTTAGTAAACACAAGATACTGGAGGTTAA 480  
 45 -----+-----+-----+-----+  
 141 D I A T K D T F G M V S K T Q D S E G K  
 160

481  
AATGTTGCTGCAACAAAGGCACTTAAAGTTAAAGATGTTGCTACATGGTTGAAGTCT 540  
5 161 N V A A T K A L K V K D V A T F G L K S  
180 541  
GGTGGAAAGCGAAGATACTGGATATGTTATTGAAATGAAAGCAGGAGCTGTAGAGGATAAG 600  
10 181 G G S E D T G Y V I E M K A G A V E D K  
200 601  
TATGGTAAAGTTGGAGATAGTACCGCAGGTATTGCAATAATCTCTCTAGTACTGGACTT 660  
15 201 Y G K V G D S T A G I A I N L P S T G L  
220 661  
GAATATGCAGGTAAAGGAACAACAATTGATTTAAATAAAAACCTTAAAGTTGATGTAACA 720  
20 221 E Y A G K G T T I D F N K T L K V D V T  
240 721  
GGTGGTTCAACACCTAGTGCTGTAGCTGTAAGTGGTTTGTAACTAAAGATGATACTGAT 780  
25 241 G G S T P S A V A V S G F V T K D D T D  
260 781  
TTAGCAAAATCAGGTACTATAATGTAAGAGTTATAATGCAAAAGAAGATCAATTGAT 840  
30 261 L A K S G T I N V R V I N A K E E S I D  
280 841  
ATAGATGCAAGCTCATATACATCAGCTGAAATTTAGCTAAAAGACATGTATGGATCCA 900  
35 281 I D A S S Y T S A E N L A K R H V F D P  
300 901  
GATGAAATTCCTGAAGCATATAAGGCAATAGTAGCATTACAAAATGATGGTATAGAGTCT 960  
40 301 D E I S E A Y K A I V A L Q N D G I E S  
320 961  
AATTTAGTTCAAGTTAGTTAATGGAAAATATCAAGTGATTTTATCCAGAAGGTAAGA 1020  
45 321 N L V Q L V N G K Y Q V I F Y P E G K R  
340

1021 TTAGAAACTAAATCAGCAAATGATAACAATAGCTAGTCAAGATAACCCAGCTAAAGTAGTT  
 1080

5	341 L E T K S A N D T I A S Q D T P A K V V
	360 *

1081  
 ATAAAGCTAATAAATTAAAAGATTAAAAGATTATGTAGATGATTAAAAACATATAAT 1140

10	361 I K A N K L K D L K D Y V D D L K T Y N
	380

1141  
 ATA TCTATTCAAATGTTGAACAGTAGCAGGAGAAGATAGAATAGAAACTGCTATAGAA 1200

15	381 N T Y S N V V T V A G E D R I E T A I E
	400

1201  
 TTAAGTAGTAAATATTATAATTCTGATGATAAAATGCAATAACTGATAAAGCAGTTAAT 1260

20	401 L S S K Y Y N S D D K N A I T D K A V N
	420

1261  
 GATATAGTATTAGTGGATCTACATCTATAGTTGATGGCTTGTCATCACCATTAGCT 1320

25	421 D I V L V G S T S I V D G L V A S P L A
	440

1321  
 TCAGAAAAAACAGCTCATTATTAACTTCAAAAGATAAATTAGATTCATCAGTAAAA 1380

30	441 S E K T A P L L L T S K D K L D S S V K
	460

1381  
 TCTGAAATAAAGAGAGTTATGAACTTAAAGAGTGACACTGGTATAAAATCTAAAGA 1440

35	461 S E I K R V M N L K S D T G I N T S K K
	480

1441  
 GTTTATTAGCTGGTGGAGTTAATTCTATCTAAAGATGTAGAAAATGAATTGAAAAC 1500

40	481 V Y L A G G V N S I S K D V E N E L K N
	500

1501  
 ATGGGTCTTAAAGTTACTAGATTATCAGGAGAAGACAGATAACGAAACTCTTAGCAATA 1560

45	501 M G L K V T R L S G E D R Y E T S L A I
	520

1561  
 GCTGATGAAATAGGTCTTGATAATGATAAAGCATTGTAGTTGGTGGTACTGGATTAGCA 1620

50	521 A D E I G L D N D K A F V V G G T G L A
	540

1621  
 GATGCTATGAGTATAGCTCCAGTTGCTCTCAACTTAAAGATGGAGATGCTACTCCAATA 1680

541 D A M S I A P V A S Q L K D G D A T P I  
560  
1681 GTAGTTGATGGAAAAGCAAAGAAATAAGTGTGATGCTAAGAGTTCTTAGGAACCT 1740  
5 V V V D G K A K E I S D D A K S F L G T  
580  
1741 TCTGATGTTGATATAATAGGTGGAAAAAATAGCGTATCTAAAGAGATTGAAGAGTCATAA 1800  
10 S D V D I I G G K N S V S K E I E E S I  
600  
1801 GATAGTGCAACTGGAAAAACTCCAGATAGAATAAGTGGAGATGATAGACAGCAACTAAT 1860  
15 D S A T G K T P D R I S G D D R Q A T N  
620  
1861 GCTGAAGTTAAAGAACGATGATTATTCACAGATGGTGAAGTTGTGAATTACTTTGTT 1920  
20 A E V L K E D D Y F T D G E V V N Y F V  
640  
1921 GCAAAGATGGTTCTACTAAAGAACGATCAATTAGTAGATGCCTTAGCAGCACCAATA 1980  
25 A K D G S T K E D Q L V D A L A A A P I  
660  
1981 GCAGGTAGATTTAAGGAGTCTCCAGCTCCAATCATACTAGCTACTGATACTTATCTTCT 2040  
30 A G R F K E S P A P I I L A T D T L S S  
680  
2041 GACCAAAATGTAGCTGTAAGTAAAGCAGTTCTAAAGATGGTGGAACTAACCTAGTTCAA 2100  
35 D Q N V A V S K A V P K D G G T N L V Q  
700  
2101 GTAGGTAAGGTATAGCTCTTCAGTTATAAACAAAATGAAAGATTATTAGATATG  
2157  
40 V G K G I A S S V I N K M K D L L D M  
719

## Appendix 8

SEQ ID No 10. Nucleotide sequence of *slpA* from *Clostridium difficile* strain 170426, PCR type 92, with translation. The putative secretory signal cleavage site ( $\Delta$ ) and site of cleavage to form the two mature SLPs ( $\blacklozenge$ ) are indicated.

1 ATGAAATAAGAAAAATATAGCAATAGCTATGTCAGGTTAACAGTTTAGCTTCGGCTGCT 60  
 10 -----+-----+-----+-----+-----+  
 1 M N K K N I A I A M S G L T V L A S A A  
 20 61 CCTTTTGTGCTGCAACTACTGGACACACAAGGTTACTGTAGTTAAAACGACTGGAAA 120  
 15 -----+-----+-----+-----+-----+  
 21 P V F A A T T G T Q G Y T V V K N D W K  
 40 Δ  
 20 121 AAAGCAGTAAACAAATTACAGGATGGACTAAAGATAATAGTATAGGAAGATAACTGTA 180  
 25 -----+-----+-----+-----+  
 41 K A V K Q L Q D G L K D N S I G K I T V  
 60 181 TCTTTTAATGATGGGTTGGGGTGAAGTAGCTCTAAAGTGCTAATAAGAAAGCGGAC 240  
 25 -----+-----+-----+-----+  
 61 S F N D G V V G E V A P K S A N K K A D  
 80 241 AGAGATGCTGCAGCTGAGAAGTTATATAATCTTGTAAACACTCAATTAGATAATTAGGT 300  
 30 -----+-----+-----+-----+  
 81 R D A A A E K L Y N L V N T Q L D K L G  
 100 301 GATGGAGATTATGTTGATTTCTGTAGATTATAATTAGAAAAAAAATAAACTAAT 360  
 35 -----+-----+-----+-----+  
 101 D G D Y V D F S V D Y N L E K K I I T N  
 120 361 CAAGCAGATGCGAGCAAGCAATTGTTACAAAGTTAAATTCACTTAATGAGAAAACCTTATT 420  
 40 -----+-----+-----+-----+  
 121 Q A D A E A I V T K L 'N S L N E K T L I  
 140 421 GATATAGCRAACTAAAGATACTTTGGAATGGTTAGTAAACACAAAGATAGTGAAGGTAAA 480  
 45 -----+-----+-----+-----+  
 + 141 D I A T K D T F G M V S K T Q D S E G K  
 50 160  
 481 AATGTTGCTGCAACAAAGGCACCTAAAGTTAAAGATGTTGCTACATTGGTTGAAGTCT 540  
 55 -----+-----+-----+-----+  
 161 N V A A T K A L K V K D V A T F G L K S  
 180

541  
 GGTGGAAGCGAAGATACTGGATATGTTGTTGAAATGAAAGCAGGAGCTGTAGAGGATAAG 600  
 -----+-----+-----+-----+-----+  
 181 G G S E D T G Y V V E M K A G A V E D K  
 5 200  
 601  
 TATGGTAAAGTTGGAGATAGTACGGCAGGTATTGCAATAAAATCTTCCTAGTAGCTGGACTT 660  
 -----+-----+-----+-----+  
 201 Y G K V G D S T A G I A I N L P S T G L  
 10 220  
 661  
 GAATATGCAGGTAAGGAACAACATTGATTAAATAAAACCTTAAAGTTGATGTAACA 720  
 -----+-----+-----+-----+  
 221 E Y A G K G T T I D F N K T L K V D V T  
 15 240  
 721  
 GGTGGTTCAACACCTAGTGCTGTAGCTGTAAGTGGTTTGTAACTAAAGATGATACTGAT 780  
 -----+-----+-----+-----+  
 241 G G S T P S A V A V S G F V T K D D T D  
 20 260  
 781  
 TTAGCAAAATCAGGTACTATAATGTAAGAGTTATAATGCAAAAGAAGAATCAATTGAT 840  
 -----+-----+-----+-----+  
 261 L A K S G T I N V R V I N A K E E S I D  
 25 280  
 841  
 ATAGATGCAAGCTCATATACATCAGCTGAAAATTAGCTAAAAGATATGTTGATCCA 900  
 -----+-----+-----+-----+  
 281 I D A S S Y T S A E N L A K R Y V F D P  
 30 300  
 901  
 GATGAAATTCTGAAGCATATAAGGCATAGTAGCATTACAAAATGATGGTATAGAGTCT 960  
 -----+-----+-----+-----+  
 301 D E I S E A Y K A I V A L Q N D G I E S  
 35 320  
 961  
 AATTTAGTCAGTTAGTTAATGGAAAATATCAAGTGATTTTTATCCAGAAGGTAAAGA 1020  
 -----+-----+-----+-----+  
 321 N L V Q L V N G K Y Q V I F Y P E G K R  
 40 340  
 1021  
 TTAGAAACTAAATCAGCAAATGATACAATAGCTAGTCAGATAACCCAGCTAAAGTAGTT 1080  
 -----+-----+-----+-----+  
 341 L E T K S A N D T I A S Q D T P A K V V  
 45 360  
 ♦  
 1081  
 ATAAAAGCTAATAAAATTAAAAGATTAAAAGATTATGTAGATGATTAAAACATATAAT 1140  
 -----+-----+-----+-----+  
 361 I K A N K L K D L K D Y V D D L K T Y N  
 380  
 1141  
 AATACTTATTCAAATGTTAACAGTAGCAGGAGAAGATAGAAATAGAAACTGCTATAGAA 1200  
 -----+-----+-----+-----+

381 N T Y S N V V T V A G E D R I E T A I E  
 400  
 1201  
 5 TTAAGTAGTAAATATTATAATTCTGATGATAAAATGCAATAACTGATAAAGCAGTTAAT 1260  
 401 L S S K Y Y N S D D D K N A I T D K A V N  
 420  
 1261  
 10 GATATAGTATTAGTTGGATCTACATCTATAGTTGATGGCTTGTCATCACATTAGCT 1320  
 421 D I V L V G S T S I V D G L V A S P L A  
 440  
 1321  
 15 TCAGAAAAAACAGCTCCATTATTATAACTTCAGAAAGATAAATTAGATTCACTCGATAAA 1380  
 441 S E K T A P L L L T S K D K L D S S V K  
 460  
 1381  
 20 TCTGAAAATAAGAGAGTTATGAACTTAAAGAGTGACACTGGTATAAATTACTCTAAAAA 1440  
 461 S E I K R V M N L K S D T G I N T S K K  
 480  
 1441  
 25 GTTATTAGCTGGTGGAGTTAATTCTATATCTAAAGATGTAGAAAATGAATTGAAAAAC 1500  
 481 V Y L A G G V N S I S K D V E N E L K N  
 500  
 1501  
 30 ATGGGTCTTAAAGTTACTAGATTATCAGGAGAAGACAGATAACGAACTTCTTAGCAATA 1560  
 501 M G L K V T R L S G E D R Y E T S L A I  
 520  
 1561  
 35 GCTGATGAAATAGGTCTTGATAATGATAAAGCATTGATGGTGGTAGCTGGATTAGCA 1620  
 521 A D E I G L D N D K A F V V G G T G L A  
 540  
 1621  
 40 GATGATGAGTATAGCTCCAGTTGCTTCTCAACTTAAAGATGGAGATGCTACTCCAATA 1680  
 541 D A M S I A P V A S Q L K D G D A T P I  
 560  
 45  
 1681  
 561 V V V D G K A K E I S D D A K S F L G T  
 580  
 1741  
 50 TCTGATGTTGATATAATAGGTGGAAAAAAATAGCGTATCTAAAGAGATTGAAGAGTCATAA 1800  
 581 S D V D I I G G K N S V S K E I E E S I  
 600  
 1801  
 55 GATAGTGCAACTGGAAAAACTCCAGATAGAATAAGTGGAGATGATAGACAAGCAACTAAT 1860

601 D S A T G K T P D R I S G D D R Q A T N  
620 1861  
5 GCTGAAGTTTAAAGAAGATGATTATTCACAGATGGTGAAGTTGTGAATTACTTTGTT 1920  
621 A E V L K E D D Y F T D G E V V N Y F V  
640 1921  
10 GCAAAAGATGGTCTACTAAAGAAGATCAATTAGTAGATGCCTTAGCAGCACCAATA 1980  
641 A K D G S T K E D Q L V D A L A A A P I  
660 1981  
15 GCAGGTAGATTTAAGGAGTCTCCAGCTCCAATCATACTAGCTACTGATACTTTATCTCT 2040  
661 A G R F K E S P A P I I L A T D T L S S  
680 2041  
20 GACCAAAATGTAGCTGTAAGTAAAGCAGTTCTAAAGATGGTGGAACTAACTTAGTTCAA 2100  
681 D Q N V A V S K A V P K D G G T N L V Q  
700 2101 2157  
25 701 V G K G I A S S V I N K M K D L L D M  
719

30

Claims

1. A vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans.
2. A vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising a *C. difficile* gene or *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof to which immunoreactivity is detected in individuals who have recovered from *C. difficile* infection.
3. A vaccine as claimed in claim 1 or 2 wherein the gene encodes a *C. difficile* surface layer protein, SlpA or variant or homologue thereof.
4. A vaccine as claimed in claim 1 or 2 wherein the peptide/polypeptide is a *C. difficile* surface layer protein, SlpA or variant or homologue thereof.
5. A vaccine as claimed in any of claims 1 to 4 wherein the vaccine comprises a chimeric nucleic acid sequence.
6. A vaccine as claimed in 5 wherein the chimeric nucleic acid sequence is derived from the 5' end of the gene, encoding the mature N-terminal moiety of SlpA from *C. difficile*.
7. A vaccine as claimed in any of claims 1 to 4 wherein the vaccine comprises a chimeric peptide/polypeptide.
8. A vaccine as claimed in 7 wherein the amino acid sequence of the chimeric peptide/polypeptide is derived from the mature N-terminal moiety of SlpA from *C. difficile*.

9. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains an amino acid sequence SEQ ID No.1 or a derivative or fragment or mutant or variant thereof.
- 5 10. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains an amino acid sequence SEQ ID No.2 or a derivative or fragment or mutant or variant thereof.
- 10 11. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.3 or a derivative or fragment or mutant or variant thereof.
- 15 12. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.4 or a derivative or fragment or mutant or variant thereof.
13. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.5 or a derivative or fragment or mutant or variant thereof.
- 20 14. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.6 or a derivative or fragment or mutant or variant thereof.
- 25 15. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.7 or a derivative or fragment or mutant or variant thereof.
16. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.8 or a derivative or fragment or mutant or variant thereof.

17. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.9 or a derivative or fragment or mutant or variant thereof.
- 5 18. A vaccine as claimed in any of claims 1 to 8 wherein the vaccine contains a nucleotide sequence SEQ ID No.10 or a derivative or fragment or mutant or variant thereof.
- 10 19. A vaccine as claimed in any preceding claim in combination with at least one other *C. difficile* sub-unit.
- 20 20. A vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising the mature N-terminal moiety of a surface layer protein, SlpA of *C. difficile* or variant or homologue thereof which is immunogenic in humans.
- 15 21. A vaccine as claimed in claim 20 wherein the N-terminal moiety of SlpA contains an amino acid sequence SEQ ID No. 1.
- 20 22. A vaccine as claimed in claim 20 wherein the N-terminal moiety of SlpA contains an amino acid sequence SEQ ID No. 2.
- 25 23. A vaccine for the treatment or prophylaxis of *C. difficile* associated disease, the vaccine comprising an immunodominant epitope derived from a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans.
24. A vaccine as claimed in any preceding claim comprising a pharmaceutically acceptable carrier.
- 30 25. A vaccine as claimed in any preceding claim in combination with a pharmacologically suitable adjuvant.

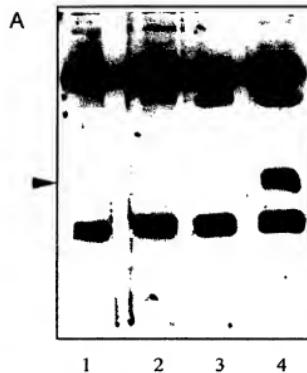
26. A vaccine as claimed in claim 25 wherein the adjuvant is interleukin 12.
27. A vaccine as claimed in claim 25 or 26 wherein the adjuvant is a heat shock protein.
- 5 28. A vaccine as claimed in any preceding claim comprising at least one other pharmaceutical product.
- 10 29. A vaccine as claimed in claim 28 wherein the pharmaceutical product is an antibiotic.
30. A vaccine as claimed in claim 29 wherein the antibiotic is selected from one or more metronidazole, amoxycillin, tetracycline or erythromycin, clarithromycin or tinidazole.
- 15 31. A vaccine as claimed in claim 28 wherein the pharmaceutical product comprises an acid-suppressing agent such as omeprazole or bismuth salts.
- 20 32. A vaccine as claimed in any preceding claim in a form for oral administration.
33. A vaccine as claimed in any preceding claim in a form for intranasal administration.
- 25 34. A vaccine as claimed in any preceding claim in a form for intravenous administration.
35. A vaccine as claimed in any preceding claim in a form for intramuscular administration.
- 30 36. A vaccine as claimed in any of claims 1 to 35 including a peptide delivery system.

37. An immunodominant epitope derived from a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof.
- 5 38. An immunodominant epitope as claimed in claim 37 wherein the *C. difficile* peptide/polypeptide contains an amino acid sequence SEQ ID No.1 or SEQ ID No.2 or a derivative or fragment or mutant or variant thereof.
- 10 39. An immunodominant epitope as claimed in claim 35 wherein the *C. difficile* peptide/polypeptide contains an amino acid sequence SEQ ID No.3 or SEQ ID No.4 or SEQ ID No.5 or SEQ ID No.6 or SEQ ID No.7 or SEQ ID No.8 or SEQ ID No. 9 or SEQ ID No. 10 or a derivative or fragment or mutant or variant thereof.
- 15 40. A chimeric nucleic acid sequence derived from the 5' end of the slpA gene encoding the mature N-terminal moiety of SlpA from *C. difficile* which is immunogenic in humans.
- 20 41. A chimeric peptide/polypeptide wherein the amino acid sequence of the chimeric peptide/polypeptide is derived from the mature N-terminal moiety of SlpA from *C. difficile*.
42. A *C. difficile* peptide comprising SEQ ID No. 1.
43. A *C. difficile* peptide comprising SEQ ID No. 2.
- 25 44. A *C. difficile* gene comprising SEQ ID No. 3.
45. A *C. difficile* gene comprising SEQ ID No. 4.
- 30 46. A *C. difficile* gene comprising SEQ ID No. 5.

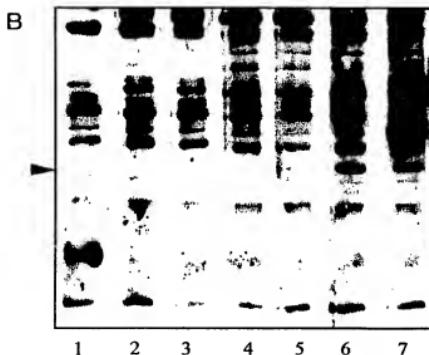
47. A *C. difficile* gene comprising SEQ ID No. 6.
48. A *C. difficile* gene comprising SEQ ID No. 7.
- 5 49. A *C. difficile* gene comprising SEQ ID No. 8.
50. A *C. difficile* gene comprising SEQ ID No. 9.
51. A *C. difficile* gene comprising SEQ ID No. 10.
- 10 52. The use of a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans in the preparation of a medicament for use in a method for the treatment or prophylaxis of *C. difficile* infection or *C. difficile* associated disease in a host.
- 15 53. The use as claimed in claim 52 wherein the medicament which is prepared is a vaccine as claimed in any of claims 1 to 36.
- 20 54. A method for preparing a vaccine for prophylaxis or treatment of *C. difficile* associated disease, the method comprising;
  - obtaining a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans; and
  - 25 forming a vaccine preparation comprised of said gene or peptide/polypeptide or derivative or fragment or mutant or variant, which is suitable for administration to a host and which when administered raises an immune response.

55. A method as claimed in claim 54 wherein the *C. difficile* peptide/polypeptide contains an amino acid sequence SEQ ID No.1 or SEQ ID No.2 or a derivative or fragment or mutant or variant thereof.
- 5 56. A method as claimed in claim 54 wherein the *C. difficile* gene contains an amino acid sequence SEQ ID No.3 or SEQ ID No.4 or SEQ ID No.5 or SEQ ID No.6 or SEQ ID No.7 or SEQ ID No.8 or SEQ ID No.9 or SEQ ID No.10 or a derivative or fragment or mutant or variant thereof.
- 10 57. A method for prophylaxis or treatment of *C. difficile* associated disease, the method comprising:
  - obtaining a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans;
  - 20 forming a vaccine preparation comprised of said gene or peptide/polypeptide or derivative or fragment or mutant or variant, and
  - 25 administering the vaccine preparation to a host to raise an immune response.
- 25 58. Monoclonal or polyclonal antibodies or fragments thereof, to a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans.
- 30 59. Monoclonal or polyclonal antibodies or fragments thereof, to *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof to which immunoreactivity is detected in individuals who have recovered from *C. difficile* infection.

60. Purified antibodies or serum obtained by immunisation of an animal with a vaccine according to any of claims 1 to 36.
- 5        61. The use of the antibodies or fragments as claimed in claims 58 and 59 in the preparation of a medicament for treatment or prophylaxis of *C. difficile* infection or *C. difficile* associated disease.
- 10      62. The use of the antibodies or serum as claimed in 60 in the preparation of a medicament for treatment or prophylaxis of *C. difficile* infection or *C. difficile* associated disease.
- 15      63. The use of the antibodies or fragments or serum as claimed in any of claims 58 to 60 for use in passive immunotherapy for established *C. difficile* infection.
64. The use of the antibodies or fragment or serum as claimed in any of claims 58 to 60 for the eradication of *C. difficile* associated disease.
- 20      65. Use of interleukin 12 as an adjuvant in *C. difficile* vaccine.
66. The use of humanised antibodies or serum for passive vaccination of an individual with *C. difficile* infection.



1 2 3 4



1 2 3 4 5 6 7

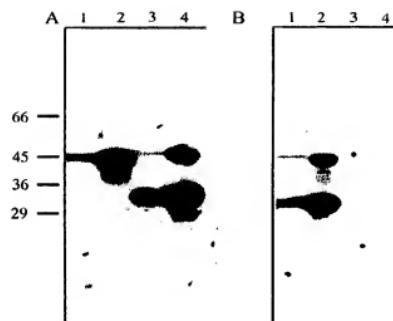


Figure 2

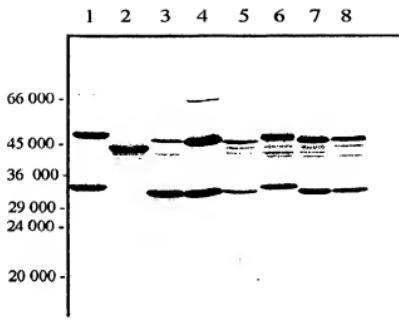


Figure 3

## SEQUENCE LISTING

<110> The Provost, Fellows & Scholars of the College of the Holy and Undivided Trinity of Queen Elizabeth, near Dublin

<120> C.difficile vaccine

<130> TRI002/C/WO

<160> 10

<170> PatentIn version 3.1

<210> 1

<211> 2157

<212> DNA

<213> Clostridium difficile

<400>	1					
atgaataaga	aaaatatagc	aatacgctatg	tcaggtttaa	cagtttttagc	ttcggctgct	60
cctgttttttg	ctgcaactac	tggAACACAA	ggttatactg	tagttaaaaaa	cgactggaaa	120
aaAGCAGTAA	ACAATTACCA	GGATGGACTA	AAAGATAATA	GTATAGGAAA	GATAACTGTA	180
tcttttaatg	atggggttgt	gggtgaagta	gctctaaaaa	gtgttaataa	gaaAGCGGAC	240
agagatgctg	cagctgagaa	gttatataat	cttggtaaca	ctcaatttga	taaatttaggt	300
gatggagatt	atgttgattt	ttctgtagat	tataatttag	aaaaaaaaat	aataactaat	360
caaggcagatg	cagaagcaat	tttttacaaag	ttaaattcac	ttaatgagaa	aacttttatt	420
gatatacgaa	ctaaagatac	tttggaaatg	tttagtaaaaa	cacaagatag	tgaaggtaaa	480
aatgttgctg	caacaaggc	actttaaagt	aaagatgttg	ctacatttgg	tttggaaatc	540
gggtggaaagcg	aagatactgg	atatgttgg	gaaatgaaag	caggagctgt	agaggataag	600
tatggtaaag	ttggagatag	tacggcagg	attgcaataa	atcttcctag	tactggactt	660
gaatatgcag	gtaaaggAAC	aacaatttgc	ttaataaaaa	cttttttttt	ttatgttaaca	720
gggttgttcaa	cacctgtgc	tgtatgttgc	agtgggtttt	taactaaaga	tgataactgt	780
ttagcaaat	caggactat	aaatgttgc	tttataatgt	aaaagaaga	atcaatttgc	840
atagatgcac	gctcatatac	atcagctgaa	aatttatgt	aaagatgtat	ttttgtatcc	900
gatgaaattt	ctgaagcata	taaggcaata	gtacgttac	aaaatgttgc	tatagatgtct	960
aattttatgc	atgtatgttgc	tttttttttt	tttatccaga	aggtaaaaga	tttttttttt	1020
tttagaaacta	aatcagccaa	tgatataata	gtctgttgc	ataccaccgc	tttttttttt	1080
ataaaagcta	ataaaattaa	agatTTTTAA	gattatgttgc	atgttttttt	tttttttttt	1140
aatacttatt	caatgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1200
ttaatgttgc	aatattataa	ttctgtatgt	aaaaatgcac	taactgtatgc	tttttttttt	1260

gatatagtat tagtggatc tacatctata gttgtggtc ttgtgcac accattagct	1320
tcagaaaaaa cagctccatt attattaact tcaaaagata aattagatc atcagtaaaa	1380
tctgaaataa agagagttat gaacttaaag agtgacactg gtataatac ttctaaaaaa	1440
gtttatTTAG ctggTggagt taattctata tctaaagatg tagaaaatga attgaaaaac	1500
atgggtctta aagttaactag attatcgga gaagacagat acgaaacttc tttagcaata	1560
gctgtgaaa taggtcttga taatgataaa gcattttag ttgggtgtac tggattagca	1620
gatgttatGA gtatagctcc agttgcttct caacttaaag atggagatgc tactccaata	1680
gtagtttagt atggaaaagc aaaagaata agtgatgatg ctaagagttt ctttaggaact	1740
tctgtgttG atataatagg tgaaaaaaat agcgtatcta aagagattga agagtcaata	1800
gatagtgcaa ctggaaaaac tccagataga ataagtggag atgatagaca agcaactaat	1860
gctgaagttt taaaagaaga tgattatttc acagatggtg aagttgtgaa ttactttgtt	1920
gcaaaagatg gttctactaa agaagatcaa ttagtagatg ctttagcagc agcacaata	1980
gcaggtatG ttaaggagtc tccagctcca atcatactag ctactgatac ttatcttct	2040
gaccaaaaatg tagctgtaaG taaaggagtt cctaaagatg gtggactaa cttagttcaa	2100
gtaggtaaag gtatagcttC ttcaGttata aacaaaatga aagatttatt agatatg	2157

<210> 2  
<211> 1830  
<212> DNA  
<213> Clostridium difficile

<400> 2 atgaaaaaaa gaaatttagc aatggctatg gcagctgtta ctgttagtagg ttctgctgtct	60
ccagtttttgcagcgttc agatgtata tcactacaag atggatcaaa tgataagttat	120
acagtatcaa atactaaagc tagtgactta gtaaaggata ttttagcagc acaaaactta	180
acaacaggtg cagttatTTT gaacaaagat acaaaagtta ctttctatga tgcaaatgag	240
aaagattctt caactccaaG tggagataaa aaagtttatt cagaacaaac tttaactaca	300
gctaatggaa atgaagatTA tgtaaagaca actttttttt attttagatgc aggagaatat	360
gctattatAG attaactta taataatgtc aaaactgttG aaattaaagt agtagcagct	420
agtggaaaaaa cagtagttgt atctagtgtat gcaaaaaata gtgcaaaaga tatagctgaa	480
aaatatgtgt tgtaagacaa agacttagaa aatgactaa aactataaa tgcctcagat	540
ttcagtaaaa ctgatagtttG ctatcaagta gttctttatc caaaaggaaa gagattacaa	600
ggtttctcaa cttatagagc tacaattat aatgaaggaa ctgcataatgg taatacacca	660
gtaatattaa ctctaaaatc tacttagtaag agtaatttaa agactgcagt agaagagttt	720

3/11

aaaaattga atgctagttt ttctaataact acaacttttag ctggtgatga cagaatacaa 780  
acagctatag agataagtaa agaatattac aataatgtat gcgagaaaatc agatcattca 840  
gctgtatgtta aagagaatgt taaaaatgtt gtatttagtag gtgcaaatgc actagtatgt 900  
ggatttagtt cggtccctt agcagcagaa aaagatgctc cactattatt aacttcaaaaa 960  
gataaatttag attcgctcagt aaaatctgaa ataaaagagag ttttagactt aaaaacttca 1020  
acagaagtaa caggaaaaac agtttatata gctggtgag ttaatagtgt atctaaagaa 1080  
gttgtaacag aattagaatc aatgggatta aaagtgaaa gatttcagg tgatgataga 1140  
tatgaaacct cttaaaaat agcaggtgaa ataggcttag ataatgataa ggcttatgt 1200  
gttggtgaa caggattagc agatgccatg agtatactt cagttgcctc tactaaattt 1260  
gatggtaatg gtgttgatg tagaacaat ggacatgcta ctccaatagt tttgtatgt 1320  
ggaaaagctg ataaaatatc tgatgactta gatagttct taggaagcgc tgatgatgt 1380  
ataatagtg gatggcaag tttatctgaa aagatggaaag aagctatatac agatgtact 1440  
ggtaaaggcg ttacaagagt taaaggcgac gatagacaag acactaactc tgaagttata 1500  
aaaacatatt atgctaatga tactgaaata gctaaagctg cagttttaga taaagattca 1560  
ggtgcttcaa gtatgtatgc aggagtatatt aatttctatg tagctaaaga tggatctaca 1620  
aaagaagatc aattatgttgc tgcatctgaa gttaggagctg ttgctggata taaaacttgc 1680  
ccagttgtat tagctactgtt ttctttatct tctgtatcaat cggttgctat aagcaaagtt 1740  
gttaggagaaa aatattctaa agatctaaca caagttggtc aaggaatagc taattcgtt 1800  
ataaaacaaaa taaaagattt attagatgt 1830

<210> 3  
<211> 2158  
<212> DNA  
<213> Clostridium difficile

```
<400> 3
atgaaaataaga aaaatatacg aatacgctatg tcaggtttaa cagttttacg ttcggctgct 60
cctgttttg ctgcaactac tggAACACAA ggttatactg tagttaaaaa cgactggaaa 120
aaaggcgttaa aacaattaca agatggacta aaagataata gtataggaaa gataactgta 180
tcttttaatg atgggggtgt gggtgaagta gctccctaaaa gtgctaataa gaaagcggac 240
agagatgctg cagctgagaa gttatataat cttgttaaca ctcattttaga taaatttagt 300
gatggagatt atgttgatt ttctgttagat tataatttag aaaacaaaat aataactaat 360
caaggcagatg cagaagcaat tgttacaaag ttaaattcac ttaatgagaa aactcttatt 420
gatatacgaa ctaaagatAC ttttggatG gtttagaaaa cacaagatAG tgaaggtaaa 480
```

```
<210> 4
<211> 2271
<212> DNA
<213> Clostridium difficile
```

<400> 4  
atgaataaga aaaatatagc aatagctatg tcaggttaa cagttttagc ttccggctgca 60  
cctgtatgg cagatgatac aaaagttga actgggtatc aaggatatac agtggtaaca 120  
agcaagtata agaaaagctgt tgaacaatta caaaaagggaa tattatgtt aagtataaca 180  
gaattaaag tttcttttga gggacttta gcatactacta taaaagttagg ttctgagctt 240  
aatgcagcag atgcaagtaa attattgttt acacaaggtag ataataaaact agataattta 300  
ggtgatggg attatgttaga ttcttaata acttctccag gtcaaggggaa taaaataact 360  
acaagtaaac ttgttcattt gaaagattta acaggtgtt cagcagatgc tataattgtt 420  
ggaacatctt cagcagatgg tgggttaca aatactggag ctgtctgtt ttctacttag 480  
acaatttcag caggaacaaa acttgcattt ttagtgcattt ttgacacagc atatacagat 540  
tcatctgaaa ctgcgggttta gattactata aaagcagata tgaatgatac taaatttgtt 600  
aaagcaggtg agacaactta ttcaactggg cttacattt aagatgggtc tacagaaaa 660  
attgttaat taggggacac tgatattata gatataacta aagctcttaa acttactgtt 720  
gttcctggaa gtaagcaac tggtaatgtt gctaaaaaa caccaggc cagtggttcaa 780  
ccagtaataa caaagcttag ataataaaat gctaaagaag aaacaataga tattgacgct 840  
agttcttagta aaacagcaca agatggatc aaaaatatg tatttataa aactgattta 900  
aatactctt ataaagtatt aaatggatc gaagcagata ctaatggatt aatagaagaa 960  
gttagtggaa aatatcaagt agttctttt ccagaaggaa aagagttt aactaagagt 1020  
gctgcaagg cttcaattgc tggatggaaat tcaccaggta aattaactct taagtcat 1080  
aagaagaaag acttaaaaga ttatgtggat gatthaagaa cataataaa tggatattca 1140  
aatgctatag aagtagcagg agaagataga atagaaactg caatagcatt aagtcaaaaa 1200  
tattataact ctgtatgtatc aatgtctata tttagatgtt cagttgatca tggatgtt 1260  
gttggaggaa atgcaatagt tggatggactt gtagttctc ctttagctc tgaaaagaaa 1320  
gctcccttat tattatctt aaaaagataaa ttagattca gctaaaggc tgaaaataaa 1380  
agagttatgtatc atataaagag tacaacaggat ataaataactt caaagaaaatgtt 1440  
ggtaggttta atcttatatac taaaagatgtt gaaaatgtt taaaagatgtt gggactttaa 1500  
gttacacatg tagcaggaga tggatgtatc gaaacttctc taaaatgc tggatgtatc 1560  
gggtttgtata atgataaaaggc atttggatgtt ggagggacac gatggcaga tgccatgtatc 1620  
atagctccatc ttgcattctca attaagaaaat gctaatggta aatggattt agctgtatgtt 1680  
gatgtcacac caatgtatgtt tggatgtatc aagctaaaaatgtt tggatgtatgtt 1740  
gatttcttag atgatttcatac agttgtatgtt ataggtggag aaaacagtgtt atctaaagat 1800

gttggaaaatg caaatagatga tgctcacaggta aaatctcccg atagatatacg tggagatgtat	1860
agacaagcaa ctaatgc当地 agttataaaa gaatcttctt attatcaaga taacttaaat	1920
aatgataaaa aagtagttaa tttctttgtat gctaaagatg gttctactaa agaagatcaa	1980
ttagtgtat cttagcagc agctccaggat gcagcaact ttgggtgtaaac tcttaattct	2040
gatggtaagc cagtagataa agatggtaaa gtattactg gttctgataa tgataaaaat	2100
aaatttagtat ctccagcacc tatagtattat gctactgatt ctttatcttc agatcaaagt	2160
gtatctataa gtaaaaggatc tgataaaagat aatggagaaa acttagttca agttggtaaa	2220
ggatatacgta cttcaggat aaaaaaaaaatg aaagatttat tagatatgtat a	2271

```
<210> 5  
<211> 2158  
<212> DNA  
<213> Clostridium difficile
```

aatacttatt caaatgttgt aacagtagca ggagaagata gaatagaaac tgctatagaa	1200
ttaagttagta aatattataa ttctgatgat aaaaatgcaa taactgataa agcagttat	1260
gatatagtat tagttggatc tacatctata gttgatggtc ttgttgcatc accattagct	1320
tcagaaaaaa cagctccatt attattagct tcaaaagata aattagattc atcagtaaaa	1380
tctgaaataa agagagttat gaacttaaag agtgacactg gtataaatac ttctaaaaaa	1440
gtttattnag ctggtgagt taattctata tctaaagatg tagaaaatga attgaaaaac	1500
atgggtctta aagttactag attatcgga gaagacatg acgaaacttc tttagcaata	1560
gctgtgaaa taggtcttga taatgataaa gcatttttag ttggtggtac tggatttagca	1620
gatgctatga gtatagctcc agttgcttct caacttaaag atggagatgc tactccaata	1680
gttagttgtatg atggaaaagc aaaagaaata agtgtatgtg ctaagagttt ctttaggaact	1740
tctgtatgtt atataatagg tggaaaaat agcgtatcta aagagattga agagtcaata	1800
gatagtgc当地 ctggaaaaac tccagataga ataagtggag atgatagaca agcaactaat	1860
gctgaagtt taaaagaaga tgattatttc acagatggtg aagttgtgaa ttacttttt	1920
gcaaaaagatg gtctactaa agaagatcaa ttagtagatg ccttagcagc agcacaata	1980
gcaggttagat ttaaggagtc tccagctcca atcatactag ctactgatac ttatcttct	2040
gaccaaaaatg tagctgtaa taaaggcgtt cttaaagatg gtggactaa cttagttcaa	2100
gttaggtaaag gtatagcttc ttcaattata aacaaaatga aagatttatt agatatgg	2158

<210> 6  
<211> 2217  
<212> DNA  
<213> Clostridium difficile

<400> 6 atgaaaataaga aggatatacg aatagctatg tcaggattaa cagtttttc ttctgcagca	60
cctgttatgg ctgttagtag ttttacagca gattataatt atactgtatg gcaaggaaaa	120
tatcaaaaag ttataactgg attacaagat ggtttaaaaa atggaaaaat aacaaatatt	180
gatgtatata ttatgtggaa ttcaatttttgg gaggttagtgc caggttctga tgctgcagct	240
gcagttacta aattaaaaag tttagttgtat gataaggtagt ataaactttagg tgatggaaaa	300
tacgttcaat ttatgtttac ttatactact aaatctataa taactaaagc agaattaaaa	360
aattattata atcaattttaga aagttttttttt aatgtatgttataa gatagaataac tttatggaaa tgaacctcaa	420
gatacaggaa ctaaaaggct tataaaaatgtt gataactgtatg gtactacttgc ttgtgcagca	480
gctgcaccat tggaaatttttcc agatataatggt atgtatgtt aacagggttgc	540
cttaaaggcag aaccaacaag taaagttttttttt aatgtatgttataa gatagaataac tttatggaaa tgaacctcaa	600

aatacaggag caactaacta tacttctgga gctgaatat ctgttccatc tacaggctta 660  
acattaaactg ctgatacaac tgcaacaaca gatgtaaata ttctgtatgt tatgagtgc 720  
tttaaattta atggtaactga tacgattagt ggattcccg ctggttcatc agcttctact 780  
cttagagcaa gtataaaaagt aataaatgc aaagaagaat ctatagatgt tgattcaagt 840  
tcacatagaa cagctgaaga tttagctgaa aaatatgtat ttaaaccaga agatgtgaat 900  
aaaacttatg aggcaactgac tgatttatataaagaaggta taacaagtaa tcttatcaact 960  
caagatggtg gaaaatatca agttgttta ttgcgtcaag gaaagagatt aactactaaa 1020  
ggagcaactg gaacttttagc agatgaaaat tctcccttta aagtaacaat aaaagcagat 1080  
aaagtaaaag acttaaaaaga ttatgttga gattaaaaaa atgctacaac tggatattca 1140  
aattctgttg ttgttagcagg tgaagataga atagaacag caatagatgt aagtagcaaa 1200  
tactataact ctgatgtatca caatgtcaata actaaagatc cagttacaaatgttta 1260  
gttggttctc aagctgtatgt tgatggcctt gtatggctc ctttagatc tggatattca 1320  
gtccctttac tattaaactt acgcaggaaaa tttagattcaatgtttaaagc tggatgttca 1380  
agagataatgg atttaaaaatc tacaacaggta gtaaatactt ctaaaaaagt ttacttagct 1440  
ggtgaggtaa actctatatac taaagatgtaa gaaaatgaat taaaagatataatggactt 1500  
gttacaagat tatcaggaga tgatagatataatgttctt tagctatagc tggatgttca 1560  
gttcttgata atgataaagc ttttgtatgtt ggaggaacag gattagcggatgtatgtat 1620  
atagctccatgttgcgttctca attaagaaac tcaatgttgc aacttgcatttaaaaggtat 1680  
gcaactccaa tagtagttgt tgatggaaaa gctaaagata taaaatgttgc agttaaaaatgg 1740  
ttcttagatgttgcgttctca attaagaaac tcaatgttgc aacttgcatttaaaaggtat 1800  
atgaaagcaatgttgcgttctca attaagaaac tcaatgttgc agttaaaaatgg 1860  
caagcaacaa atgctaaatgttgcgttctca attaagaaac tcaatgttgc agttaaaaatgg 1920  
aacttctttgttgcgttctca attaagaaac tcaatgttgc agttaaaaatgg 1980  
ggtgctgcatttgcgttctca attaagaaac tcaatgttgc agttaaaaatgg 2040  
gtctgataaaaatgttgcgttctca attaagaaac tcaatgttgc agttaaaaatgg 2100  
caaaaatgttgcgttctca attaagaaac tcaatgttgc agttaaaaatgg 2160  
gttggtaaaatgttgcgttctca attaagaaac tcaatgttgc agttaaaaatgg 2217

<210> 7  
<211> 2145  
<212> DNA  
<213> Clostridium difficile

<400> 7  
atgaaaataaga aaaacttagc aatggctatg gcagcagttt ctgttgtgg ttctgcagcg 60  
ccaaatatttgc cagatgtac tacgccaggatatactgtat tgaaaaatga ttggaaaaaa 120  
gcagtaaaac aattacaaga tgggtgaaa aataaaacta tatcaacaat aaagggtgtct 180  
tttaatggaa actctgttg agaagttaca ccagccaggatctggagcaaa aaaagcagat 240  
agagatgctg cagctgaaa gttatataat ttagtaata cacaattaga taaacttagt 300  
gatggagatt acgttactt tgaagtaact tataatttagt ctactcaaattt aattacaaaa 360  
gcagaaggcaggcaggatct tacaaaatttata caacaatata atgataaaagt acttataat 420  
tctgcacacg atacagtaaa agttagtggatctgtatcacac aagttgtatg caaaaatgtt 480  
gcagctaacc cacttaaagt tagtgatatg tatacaatac catctgttat tactggaagt 540  
gatgattctg ggtatagtat tgcataacca acagaaaaga ctacaaggatctt attgtatgg 600  
acgggtggatctg atgcaactgc aggttaaagca ataacagtagt atacagcttc aatgaagct 660  
tttgctggaa atggaaaggatattgactac aataaatcat tcaaaggacac tgcatacaggaa 720  
gatgaaacag ttaagacaag cgggttgta cttaaagatg caagtgtat ggtgcacaca 780  
ggtaactataa aagtttagt tacaagtgc aaaaagaat ctattgtatg ggttcaagt 840  
tcatatatta gtgtgtggaaa tttagctaaa aataatgttat ttaatcttca agaggittct 900  
gaagttata atgcaataatgt tgcattacaa aatgtggaa tagaattctgtat ttagtacaa 960  
tttagttaatg gaaaatatca agttatccc tatccagaag gaaaaagatt agaaactaaa 1020  
tctgcagata taatagctgtatccagatgtt ccagctaaa taactataaa agctaataaa 1080  
ttaaaagatt taaaagatta tgcattatgtat ttaaaaatcat acaataatac ttactcaaatt 1140  
gtgttaacag tagcaggaga agatagaata gaaactgtca tagaattaaatg tagtaatataat 1200  
tataattctg atgataaaaaa tgcaataact gatgtgcag ttaataatataat agtatttagtt 1260  
ggatctcatat ctatgtgtatccat gatgtgcag ttaataatataat agtatttagtt 1320  
ccattattat taacttcaaa agataaatttattt gattcatcg taaaatctgtat gataaaaaga 1380  
gttatgtacttataaagatgtatccat gatgtgcag ttaataatataat agtatttagtt 1440  
ggatctcatat ctatgtgtatccat gatgtgcag ttaataatataat agtatttagtt 1500  
acttagatttgc caggagaaga cagatacggaa acttctttagt caatgtgc tggatgtat 1560  
cttgataatgtatccat gatgtgcag ttaataatataat agtatttagtt 1620  
gctccaggatccat taaagatggatccat gatgtgc tggatgtat 1680  
aaagcaaaatgtatccat gatgtgc tggatgtat 1740  
atagttggatccat gatgtgc tggatgtat 1800

aaaactccag atagaataag tggagatgac agacaagcaa ctaatgctga agttttaaaa	1860
gaagatgatt attcaaaga tggtaagtt gtgaattact ttgttgc当地 agatggctc	1920
actaaagaag atcaattagt agatgc当地 cagc当地 acatgc当地 tagattnaag	1980
gagtc当地 ctc当地 atcaattcat actagctact gatactttat ct当地 gagcca aatgttagct	2040
gtaagtaaag cagtc当地 aaagatggta actaacttag tt当地 agttagg taaaggtaa	2100
gcttcttc当地 ttataaaca aatgaaagat ttattagata tgtaa	2145

<210> 8  
<211> 2158  
<212> DNA  
<213> Clostridium difficile

<400> 8 atgaaataaga aaaatatacg aatagctatg tc当地 gtttgc当地 cagttttagc tt当地 ggctgct	60
cctgttttg ct当地 caactac tggacacaa gg当地 tatactg tagttaaaa cgactggaaa	120
aaagc当地 taa aacaattaca agatggacta aaagataata gtataggaaa gataactgt	180
tctttaatg atggggttgc当地 gggtaagta gctcttaaaa gt当地 taataaa gaaaggc当地	240
agagatgctg cagctgagaa gt当地 tataat cttgttaaca ct当地 atttgc当地 taaatttaggt	300
gatggagatt atgttgc当地 tt当地 ttagat tataatttag aaaaaaaaaat ataactaat	360
caaggc当地 tggatc当地 cagaagc当地 ct当地 ttttgc当地 tt当地 aatttgc当地 ttaatgagaa aacttttatt	420
gatatacgaa ctaaaagatac ttttgc当地 tt当地 aatgatgaa cacaagatag tgaaggtaaa	480
aatgttgc当地 caacaaggc actttaaagg taaatgttgc当地 ct当地 atttgc当地 tt当地 aatgttgc当地	540
gg当地 ggatc当地 aagatactgg atatgttatt gaaatgaaag caggactgt agaggataag	600
tatggtaaag tt当地 gagatg tacggc当地 atttgc当地 atcttgc当地 tactggactt	660
gaatatgc当地 gagatgaaac aacaattgtat ttaataaa ct当地 taaaatgt tgatgt	720
gg当地 ggatc当地 cacctgtgc ttagatgtc当地 agtgggttgc当地 taactaaaga tgatactgt	780
tttagcaaaat caggacttat aatgttgc当地 gttataatg caaaaggaa atcaattgtat	840
atagatgcaat gctcatatac atcagatgtc当地 aatgttgc当地 aagacatgt atttgtatcc	900
gatgaaattt ct当地 aaggc当地 taaggc当地 atcagatgtc当地 aatgttgc当地 tatagatgtct	960
aatgttgc当地 agttagttaa tgaaaatgtat caatgttgc当地 tt当地 tccaga aggtaaaaga	1020
tttagaaacta aatcagc当地 aaatgttgc当地 gttataatg ctagatgtc当地 atacaccgc当地 taaatgtat	1080
ataaaaggc当地 ataaatgtat agattnaag gattatgtat atgattnaag aacatataat	1140
aatacttatt caatgttgc当地 aacatgttgc当地 ggatc当地 gagatgaaac tgatgtat	1200
ttaatgtat aatattataa ttctgtat aatgttgc当地 taactgtat aagcgtat	1260

gatatagtat	tagttggatc	tacatctata	gttgatggtc	ttgttgcattc	accatttagct	1320
tcagaaaaaa	cagctccatt	attattaact	tcaaaagata	aattagattc	atcagtaaaa	1380
tctgaaataa	agagagttat	gaacttaaag	agtgcacatg	gtataaatac	ttctaaaaaa	1440
gttattttag	ctgggtggagt	taattctata	tctaaagatg	tagaaaatga	attgaaaaac	1500
atgggtctta	aagttactag	attatcgga	gaagacatg	acgaaacttc	tttagcaata	1560
gctgatgaaa	taggtcttga	taatgataaa	gcatttttag	ttgggtgtac	tggattagca	1620
gatgctatga	gtatagctcc	agttgcttct	caacttaaag	atggagatgc	tactccaata	1680
gtatgttag	atggaaaagc	aaaagaata	agtgtatgt	ctaagagttt	cttaggaact	1740
tctgatgtt	ataataatagg	tggaaaaat	agcgtatcta	aagagattga	agagtcaata	1800
gatagtgcaa	ctggaaaac	tccagataga	ataagtggag	atgatagaca	agcaactaat	1860
gctgaagttt	taaaagaaga	tgattatttc	acagatggt	aagttgtgaa	ttactttgtt	1920
gcaaaagatg	gttctactaa	agaagatcaa	ttagtagatg	ccttagcagc	agcaccaata	1980
gcaggttagat	ttaaggagtc	tccagctcca	atcatactag	ctactgatac	tttatcttct	2040
gacccaaaatg	tagctgtaaag	taaaggcagtt	cctaaagatg	gtggactaa	cttagttcaa	2100
gtaggtaaag	gtatagcttc	ttcagttata	aacaaaatga	aagatttatt	agatatga	2158

&lt;210&gt; 9

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Clostridium difficile

&lt;400&gt; 9

Asp	Lys	Thr	Lys	Val	Glu	Thr	Ala	Asp	Gln	Gly	Tyr	Thr	Val	Val	Gln
1				5					10					15	

Ser	Lys	Tyr	Lys
	20		

&lt;210&gt; 10

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Clostridium difficile

&lt;400&gt; 10

Ala	Thr	Thr	Gly	Thr	Gln	Gly	Tyr	Thr	Val	Val	Lys	Asn	Asp	Gly	Lys
1				5					10				15		

Lys	Ala	Val	Lys
	20		

SEQ ID No. 1  
(Strain 171500)

DKTKVETADQGYTVVQSKYK

SEQ ID No. 2  
(Strain 170324)

ATTTGTQGYTVVKNDGKKAVK

SEQ ID No. 3  
(Strain 171500 DNA)

ATGAATAAGAAAAATAGCAATAGCTATGTCAGGTTAACAGTTTAGCTTCGGCTGCACCT  
GTATTGTCAGATGATACAAGAAGTGAACACTGGTGATCAAGGATATAACAGTGGTACAAGCAA  
GTATAAGAAAGCTGTGAAACATTACAAGAAATTAGATGGAAGTATAACAGAAATT  
AAGTTTCTTGAGGGAACTTAGCATCTACTATAAAAGTAGGTTCTGAGCTTAATGCAGCAG  
ATGCAAGTAAATTAGTTTACAAAGTAGATAATAAAACTAGATAATTAGGTGATGGAGATT  
ATGAGATTCTTAAATCTTCAGGTCAGGGATAAAATAACTACAAGTAACATTGTTG  
CATTGAAGATTACAGGTGCTTCAGCAGATGCTATAATTGCTGGAACATCTTCAGCAGATG  
GTGTTGTCAGATCTGGAGCTGCTAGTGGTTCTACTGAGACAAATTCAAGCAGGAACAAAAC  
TTGCAATGTCAGCTATTGTCACAGCATATACAGATTCTGAAACTCGCGTTAAGGATTA  
CTATAAAAGCAGATATGAATGATACTAAATTGTTAAAGCAGGTGAGACAACTTATTCAACTG  
GGCTTACATTGAAAGTGGTCTCACAAAAAAATTGTTAAATTAGGGGACAGTGATATTAG  
ATATAACTAAAGCTTAAACTTACTGTTGTTCTGGAAAGTAAAGCAACTGTTAACGTTGCTG  
AAAAAACACCAAGTGCAGTGTCAACCGATAAAACAAAGCTGAAATAAAATGCTAA  
GAAGAAAGATAATGTTGACGCTAGTTCTAGTAAACAGCACAGATTAGCTAAAAAATA  
TGTATTAAATAAAACTGTTAAACTCTTAAATGATTAAATGGAGATGAGCAGATAC  
TAATGGATAATAGAAGAAGTGTAGTGGAAATATCAAGTGTGTTTATCCAGAAGGAAAAA  
GAGTTCAACTAAAGGTGCTCAAGGGCTCAATTGCTGATGAAAATTCAACCGATTAAATTAA  
CTCTTAAGTCAGATAAGAAGAAAGCTTAAAGATTATGTGGATGATTAAAGAACATATAAT  
AATGGATTACAAATGCTATAGAAGTAGCAGGAGAAGTAGAATAGAAACTGCAATGACATT  
AAGTCAAAATATTATAACTCTGATGATGAAAATGCTATTAGAGATTCAAGTGTATAATGT  
AGTATTGTTGGAGGAATGCAATAGTTGATGGACTTGTAGCTTCCCTTATAGCTCTGAAAAA  
GAAAGCTCTTATTATTCCTAAAGATAATTAGTCAACGCTAAAGCAGTAAAGATGCTACACC  
GAGATTGATGATATAAAAGAGTACAACAGTTAAACTACTTCAAAAGAAAGTTTATTAGCTG  
GTGGAGTTATTCTATCTAAAGAAGTAGAAAAGTAATTAAAAGATATGGGACTTAAAGTT  
ACAAGGATTAGCAGGGAGATGATGATGAAACTCTCTAAAATAGCTGATGATGAAAGTAGGCT  
TGATAATGATAAAAGCATTTGAGTTGGAGGAACAGGATTAGCAGATGCCATGAGTATAGCTC  
AGTTGCACTCAATTAAAGAAATGCTAAATGGTAAAGCTTATAGCTGATGTTAGCTGATGCTACACC  
AATAGTAGTTGAGATGGAAAAGCTAAAACCTAAATGATGATGTTAGCTGATGTTAGCTTATGATG  
ATTCAACAGTTGATATAATAGTGGAGAAAACAGTGTATCTAAAGATGTTGAAAATGCAATA  
GATGATGCTACAGGTAATCTCCAGATGATGATGTTGGAGATGATGACAAGCAACTAATGC  
AAAAGTTAAAGAAACTCTTATTATCAAGATACTTAAATAATGATAAAAAGTAGTTAA  
TTTCTTGTAGCTAAAGATGGTTCTACTAAAGAAGATCAATTAGTGTGATGCTTATGCGAGCAGCT  
CCAGTTGCAAGAAACTTGTGTAACCTTAACTCTGATGTTGAAAGCCAGTAGATAAAAGATGGT  
AAAGtATAACTGGTTCTGATAATGATAAAAATAATTAGTATCTCCAGCACCTATAGTATTAG  
CTACTGATTCTTATCTCAGATCaAAGTGTATCTATAAGTAaAGTCTGATAAAGATAATGG  
AGAAAACCTAGTTCAAGTTGTTAAAGGTATAGCTACTTCAGTTATAACAAAATGAAAGATIT  
ATTAGATATGTA

SEQ ID No. 4  
(Strain 172450 DNA)

ATGAAAAAAAAGAAATTAGCAATGGCTATGGCAGCTGTTACTGTAGTAGGTTCTGCTGTC  
CTTGTGAGCAGCTTACAGATGAATATCACTACAAGATGGTACAATGATAAGTATACAGTA  
TCAAATACTAAAGCTAGTACTTAGTAAGGATATTAGCAGCACAAAACCTAACACAGGT  
GCAGTTTTGAACAAAGATACAAAAGTACTTTCTAGAACAACCTTAACTACAGCTAATGGAA  
ACTTCAACTGGAGATAAAAAAGTTTACAGAACAACTTAACTACAGCTAATGGAA  
AGATTATGTAAGACAACCTTTAAAAAATTAGTGCAGGAGAAATAGCTTATTAGATTAAAC  
TTATAATGCTAAACACTGTTGAAATTAAAGTAGTGCAGCTAGTGTGAAAAAAACAGTAGTGT  
ATCTAGTGTGCAGAAAATAGTGCAGAACAGATATACTGTAAGGAAATATGTGTTGAAGACAAAG  
ACTTAGAAAATAGCACTAAAACATATAATGCTCAGATTTCAGTAAACTGATAGTTACTAC  
AAGTAGTTTATCAGAAAAGGAAAGATTACAAGGTTCTCACTTATAGAGCTACAAATT  
ATAATGAAGGAACGTGATATGTAATACACAGTAATATAACTCTAAACTACTAGTAAGA  
GTAATTAAAGACTGCGAGTAAAGGATTTACAAAATGTTAGTGCAGTTTACTACAA  
CTTAGCTGGTGTGACAGAATACAAACAGCTATAGAGATAAGTAAGAATATAACAAAT  
GATGGCGAGAAATCAGATCATTAGCTGATGTTAAAGAGAAATGTTAAAATGTTGATTAGTA  
GGTGCAGAACTGCACTAGTGTGGATTAGTGTGCGCCCTTGTAGCAGCAGGAAAAGATGCT  
CTTATTATAACTCAGAAATGAGTGTGCACTAAATCTGAAATAAGAGAGTTG  
GACTTAAAACCTCAACAGAAGTAACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAATAG  
TGTATCTAAAGGAGTGTGAAACAGAATTAAGCATGGGATTAAGGTTAAAGGTTAGATTCTCAG  
GTGATGATAGATAAGAAACTCTTTAAAATCAGGCTGAAATAGGCTTAGATAATGATAAG  
GCTTATGTTAGTGTGGAAACAGGATTAGCAGATGCCATGAGTATAGCTTGTCTACT  
AAATTAGTGTGAAATGTTGTTGAGTAGAGACAAATGGACATGCTACTCCAATAGTTGTTG  
GATGAAAAGCTGTATAAATATCTGATGACTTAGTGTAGTTCTAGGAAGCGCTGATGAGAT  
ATAATAGTGGATTGCAAGTGATCTGAAAGATGAGAAGCTTATAGCTGCTGGT  
AAAGGCGTTCAACAGGTTAAAGGCAGATGACAAAGACACTAATGTAAGTTAAAAC  
ATATTATGCTAATGATGACTGAAATAGCTAAAGCTGAGTTTAGATAAAAGATTCAAGGCT  
AAGTAGTGTGAGGAGATTAAATTCTATGCTGAGTCAAAAGATGATCTACAAAAGAGATCA  
ATTAGTTGATGCAATTAGCAGTAGGAGCTGTTGAGTATAACTGCTCAGTTGTTAGT  
ACTGATTCTTATCTGATCAATCGGTGCTATAAGCAAAGTTGAGGAGAAAATATTCTA  
AAGATTAAACACAAGTGGTCAAGGAATAGCTAATTCAGTTATAAAACAAAATGAAAGATTAT  
TAGATATG

SEQ ID No. 5  
(Strain 170324 DNA)

ATGAATAAGAAAATATGCAATAGCTATGTCAGGTTAACAGTTTAGCTCGGCTGCTCC  
GTTTTCTGCAACTACTGGAACACAAGTTACTGTAGTTAAAACGACTGGAAAAAGCA  
GTAACAAATACAAAGATGGACTAAAGGATAATAGTGTAGGAAAGGATAACTGTATCTTAA  
TGATGGGTTGGTGTGAGTGTCTCTAAAGTGTCTAAAGAAGCGGACAGAGATGCTG  
CAGCTGAGAGATTTATAATCTGTTAACACTCAATTAGATAAAATTAGGTGATGGAGATTATG  
TTGATTTCTGTAGATTATAATTAGAAAACAAAATAACTATCAAGCAGATGCGAGAAG  
CAATTGTTACAAAGTAAATTCACTTAAATGAGAAAACCTTATGATAGCAACTAAAGATA  
CTTTGGATGGTTAGTAAACACAAGATAGTGAAGGTTAAAGTGTGCAACAAAGGCA  
CTTAAAGTTAAAGATGTTGCTACATTGGTTGAAGTGTGGGAGCGAAGAGTACTGGATAT  
GTGTTGAAATGAAAGCAGGAGCTGTAGAGGATAATGTTAGTGTAAAGTTGGAGATACTACGGC  
AGGTATTGCAATAAAATCTCTAGTACTGGCTTGAATATGCAAGGTTAACACCTAGTGTGCTGAGTGA  
TTTAATAAAACTTTAAAAGTTGATGTAACAGGTGGTCAACACTAGTGTGCTGAGTGA  
TGGTTGTTAACTAAAGATGATACTGATTAGCAGGTTAACACTAGGTACTATAATGTAAGAGTT  
AAATGCAAAAGAAGAATCAATTGATATAGTGTCAAGCTCATACATCAGCTGAAATTAG  
CTAAAAGATGATGTTGATGCCAGATGAAATTCTGAGCTATAAGGCAATAGTAGCATTAC  
AAAATGATGTTAGAGTCTAACTTAGTGTGTTAGTGTAAATGGAAAATATCAAGTGT  
ATCCAGAAGGTTAAAGGATTAGAAACAAATCAGCAATGATCAATAGCTAGTCAAGATGAA  
CCAGTAAAGTGTAAAGGCTAATGAAATTAAAAGATTAGTTAAAGATTAGTGTGATGATT  
AAAACATATAATATCAATTGTAACAGTAGCAGGAGAAGATAGAATAGAAAC

TGCTATAGAATTAAAGTAGTAAATATAATCTGATGATAAAAATGCAATAACTGATAAACG  
AGTTAATGATATAGTATTAGTGGATCTACATCTATAGTTGATGGCTTGTCATCACCATTA  
GCTTCAGAAAAAAACAGCTCATTATATTAAACTCTAAAAGATAAATTAGATTCATCGTAAA  
TCTGAAGAAAAGAGAGCTTAAAGAGTGAACCTTAAAGAGTGACACTGGTATAAATCTTAAAAGG  
TATTAGTGTGGAGGTTAATTCTATATCTAAAGATGTAGAAAAGAATGAAATTGAAAACATGGT  
CTAAAGTTACTAGATTACAGGAGAAGACAGATCGAAAATCTTACAGCAATAGCTGATGAA  
ATAGGTCTTGTATAATGATAAAAGCATTITGTAGTTGGTACTGGATTAGCAGATGCTATGAGT  
ATAGCTCCAGTGTCTCAACTTAAAGATGGAGATGCTACTCCAATAGTAGTTGATAGTGG  
AAAGCAAAGAAATAAGTGTAGTCTAAAGAGTTGAAGACTTCTAGGAACCTCTGATGTGATAATA  
GGTGGAAAAAAATAGCGTATCTAAAGAGATGGAGACTGAACTAGATAGTGCACCTGAAAAC  
TCCAGATAGATAAATGAGGAGATGATGACAAGCAACTATGCTGAAGTTAAAGAAGATG  
ATTATTTCACAGATGGAAAGTGTGAAATTCTTGTGAAAGTGGTTCTAGGAAATCTTAAAGAAG  
ATCAATAGTAGATGCCCTAGCAGCACCAATAGCAGTAGATTAAAGGAGTCTCAGCTC  
CAATCATACTAGTACTGATCTTACCTCTGACCAAAATGAGTGTAGTAAAGCAGITC  
CTAAAGATGTTGAACTAATCTAGTCAAGTAGGTTAGGTTAGCTCAGTTAAACAA  
AAATGAAAGATTATTAGATATGG

SEQ ID No. 6

(Strain 171448 DNA)

ATGAATAAGAAAAAAATATGCAATAGCTATGTCAGGTTAACAGTTTAGCTCGGCTGCTCC  
GTTTTGCTGCAACTACTGGAACACAAGTTACTGTAGTTAAAACGACTGGAAAAAGCA  
GTAACACATTACAAGATGGACTAAAGATAATAGTATAGGAAAGATAACTGTATCTTAA  
TGATGGGGTGTGGGTAAGTAGCTCTTAAAGTGTCTAAAGGAAAGCGGACAGAGATGCTG  
CAGCTGAGAAGTTATATACTTGTAAACACTCAATTAGATAAATTAGGTGATGGAGATTATG  
TTGATTTCTGTAGATTATAATTAGAAAACAAAATAACTAATCAAGCAGATGCGAGAAG  
CAATTGTTACAAAGTTAAATTCACTTAACTGAGAAAACACTTATGTGATATAGCAACTAAAGATA  
CTTTGGAAATGGTTAGTAAACACAAGATGTTGGAGGTTAAAATGTTGCTGCAACAAAGGCA  
CTTAAAGTTAAAGATGTTGCTACATTGGTTGAAGTCTGGGAAGCGGAAGATACTGGAT  
GTTGTTGAAATGAAAAGCAGGAGCTGTAGAGGATAATGTTAAAGTTGGAGATAGTACGGC  
AGGTATTGCAATAATCTCTAGTACTGGACTITGAATATGCAAGGAAAGGAAACAAATTG  
TTTAATAAAACCTTAAAGTGTGATGAAACAGGTTCACACCTAGTGTGTTAGCTGTAAG  
TGGTTTGTAACTTAAAGGATGACTGTTAGCAAAACTCAGGACTATATAATGTAAGAGTT  
AAATGCAAAAGAAAGAATCAATTGATATAGTGCAGCTCATATACATCAGCTGAAATTAG  
CTAAAAGATGATGTTGATGCTACATTGGATCCAGATGAAATTCTCTGAGCATATAAGGCAATAGTCATTAC  
AAAATGATGTTGATGACTCTAATTGTTGCTAGTGTGTTAATGGAAAATATAAGTGTGATT  
ATCCAGAAAGTTAAAGATTAGAAACTAAATCAGCAATGATACAATAGCTAGTCAAGATA  
CCAGCTAAAGTACTTAAAGCTAAATTAAAGATTTAAAGATTATGATGATGTT  
AAAACATATAATACATTATCTCAAACTTGTAAACAGTAGCAGGAGAAAGATAAGAAAC  
TGCTATAGAATTAAAGTAGTAAATTATAATCTGATGATAAAAATGCAATAACTGATAAACG  
AGTTAATGATATAGTATTAGTGTGGATCATCTATAGTTGATGTTCTGTCATCACCATTA  
GCTTCAGAAAAAAACAGCTCATTATATTAGCTCAAAAGATAAATTAGATTCTACAGTAAA  
TCTGAAATAAAAGAGAGTTGAACTTAAAGAGTGAACACTGGTATAAATACTCTTAAAAGG  
TATTAGTGTGGAGGTTAATTCTATATCTAAAGATGTAGAAAATGAAATTGAAAACATGGG  
CTTAAAGTTACTAGATTACAGGAGAAAGACAGATCGAAAATCTTACAGCAATAGCTGATGAA  
ATAGGTCTTGTATAATGATAAAAGCATTITGTAGTTGGTACTGGATTAGCAGATGCTATGAGT  
ATAGCTCCAGTGTCTCAACTTAAAGATGGAGATGCTACTCCAATAGTAGTTGATAGTGG  
AAAGCAAAGAAATAAGTGTAGTCTAAAGAGATGGAGACTGAACTAGTGTGAGTTAAAGAAGATG  
GGTGGAAAAAAATAGCGTATCTAAAGAGATGGAGACTGAACTAGTGTGATAATA  
TCCAGATAGATAAATGAGGAGATGATGACAAGCAACTATGCTGAAGTTAAAGGAGTCTCAGCTC  
CAATCATACTAGTACTGATCTTACCTCTGACCAAAATGAGTGTAGTAAAGCAGITC  
CTAAAGATGTTGAACTAATCTAGTCAAGTAGGTTAGGTTAGCTCAGTTAAACAA  
AAATGAAAGATTATTAGATATGG

## SEQ ID No. 7

(Strain 171862 DNA)

ATGAATAAGAAAAACTTAGCAATGGCATGGCAGCAGTTACTGTTGTGGGTTCTGCAGCGCCA  
 ATATTCGAGATAGTACTACGCCAGGTTAATCTGTAAGTGAAGAAATGATGGAAAAAAGCAGT  
 AAAACAATTACAAGATGGGTTGAAAATAAAACATAACAAATAAGGTGTCCTTTAATG  
 GAAACTCTGTGGAGAACGTTACACCGCCAGTCTGGAGCAAAAAAAGCAGATAGAGATGCT  
 GCAGCTGAAAAGTTATAATTAGTAAACACAATTAGATAAACAGTTGATGGAGATTAC  
 GTTCACTTGAAGTAACCTATAATTAGTACTCTAACAAATAATTACAAAGCAGAGGCA  
 GTTCTTACAAAATTACACAAATTAGATAAAAGTACTTATAAAATTCTGCAACAGATACAGTA  
 AAAGGTAGGGTACTGATACACAAGTGTAGCAAAATGTTGAGCTAACCCACCTAAAGTT  
 AGTATGATGATGATACACCTGCTTACTGGAGTGTGATCTGGGTATAGTATTGCT  
 AAACCAACAGAAAAGACTAACAGTTATTGATGGTACGGTGTGATGCAACTCGAGGTTAAA  
 GCAATAACAGTAGATAACAGCTCAATGAAGCCTTGTGGAAATGGAAAGGTTATTGACTAC  
 AATAAACTATTCAAAGAACGACTGCAACAGGAGAACGAGCTTAAGACAAGCGGGGTTGACT  
 TAAAGATGCAAGTGTATGCTGCAACAGGTTACTATAAAAGTTAGAGTACAGTGCAGGAAAG  
 AAGAATCTATTGATGGATTCAAGTCTCATATATTAGTGTGAAAATTTAGCTAAAAAATATG  
 TATTAACTCTAAAGGGTTCTGAGCTTAAATGCAATAGTTGCAATTACAAAAATGATGGAA  
 TAGAATCTGATTAGTACAAATTAGTTAATGGAAATATCAAGTATTGTTTCTATCCAGAAAGGAA  
 AAAGATTAGAAACTAAATCTGAGGATATAATAGCTGTGAGCTAGATAGTCCAGCTAAATAACT  
 ATAAAGCTAATAAAATTAAGGTTAAAGGTTAAAGGTTATTGATAGTGTGATTAAAAACATAACATAA  
 TACTTACTCAATGTGTAACAGTAGCAGGAGAAGATAGAATAGAAACTGCTATAGAATTAA  
 GTAGTAAATATTATAATCTGATGATAAAAAGCATAACTGATGATGAGCTTAAATATAG  
 TATTAGTGGATCATCTATAGTTGATGGTCTGTGTCATCACCATTAGCTTCAAGAAAAC  
 AGCTCATTATTATAACTCAAAGATAATTAGATTCATCAGTAAATCTGAGATAAAAAG  
 AGTTATGAACTTAAAGAGTGACTCTGGTATAAATACTCTAAAAAAAGTTTATTAGCTGGTGG  
 AGTTAATCTATATCTAAAGATGTAGAAGATGAAATTGAAAATATGGGCTTAAAGTTACTAG  
 ATTACAGGAGAACGAGATACGAAACTCTTGTGCAATAGCTGTGATGAAATAGGCTTGTGATAA  
 TGATGAGGAACTTGTGAGTTGGTGTACTGGATTGGCAGAGTCTGATGATGTTAGCTCCAGTGTG  
 TTCTCAACTTAAAGATGGAGATGCTACTCAAATAGTTAGTGTGATGATAATAGGTGGAAAAAAT  
 TAAGTGTGATGCTAAAGGTTCTGATGTTAGGAACTCTGTGATGTTAGTATAATAGGTGGAAAAAAT  
 GCGTATCTAAAGAGGAGAAGTCAATAGTGTGCAACTGGAAAATCAGGAGATAGAATA  
 AGTGGAGATGACAGAACAGCAACTAATGCTGAAGTTAAAAGAAGATGATTATTCAAGA  
 TGGTAAGTTGTGAATTACTTGTGCAAAAGATGTTCTACTAAAGAAGATCAGGTTAGTGA  
 TGCATTAGCAGCAGCACCAATAGCAGGTAGATTAAAGGATCTCCAGCTCAATCATACTAGC  
 TACTGATACTTATCTCTGACCAAATAGCTGTAAGTAAAGCAGTCTCTAAAGATGGTGG  
 AACTAACCTAGTCAAGTAGGTAAAGGTATAGCTTCTCAGTTATAACAAACAAATGAAAGATT  
 ATTAGATATGTTA

## SEQ ID No. 8

(Strain 173644 DNA)

ATGAATAAGAAGGATATAGCAATAGCTATGTCAGGATTAACAGTATTAGCTTCTGCAGCACCT  
 GTATTGCTGTAGTGTGTTACAGCAGATTATAATTACTGTAAGTGCAGGAAATATCAA  
 AAAGTTAACTGGATTACAAGATGGTTAAAAATGGAAAATAACAAATATTGATGTAAT  
 ATTGATGGAAGTCAATTGGTGAAGGTAGTGCCTGGAGCTGAGCTGATGCTGAGCTGACTAA  
 ATTAAAAAGTTAGTTGATGATAAGTAACTTAACTGTTAGTGTGATGAAAATACGTTCAATTAA  
 TGTTACTTAACTACTAAATCTATAATACTAAAGCAGAATTAAAAAAATTATTAATCAATT  
 AGAAAAGTAAAGATGAAACTTATAGGAAATGAACCTCAAGATACAGGAACCTAAAGGTC  
 TTATAAAAGCTGATCTGTTAGTACTGCTGTGTCAGCAGCTGCACCATGAAATTATCAG  
 ATATATTAAGCTTATGTTAGTGTGAAAGTAAACAGGTTACTTAAAGCAGAACCAAACAGTAAAG  
 TAAGCGCTGGTAAAGTCAAGGTCTAAAATGGAAATCAGGAGCAACTAATCTACTCTG  
 GAGCTGAAATATCTGTCTCAACTACAGGCTTAACTTAAGTGTGACTGATACGATTAGTGGATT  
 ATGTTAAATTTCTGTTAGTGTGCTTAAATTAAGTGTGACTGATACGATTAGTGGATT  
 CCCAGCTGGTCTCATCAGCTTCAACTCTAGAGCAAGTAAAGTATAATGCAAAAGAAGA  
 ATCTATAGATGTTGATTCAAGTTCACATAGAACAGCTGAAGATTTAGCTGAAAAATATGTTATT

TAACCCAGAAGAGTGAATAAAAACTTATGAGGCACITGACTGATTITATATAAAAGAACGGTATAA  
 CAAGTAATCTTACTCAAGATGGGAAAATATCAAGTTGTTTATITGCTCAAGGAAAGA  
 GATTAACACTAAAGGAGCACTGGAACTTTAGCAGATGAAAATTCTCCTTAAAGTAACAA  
 TAAAAGCAGATAAGTAAAAGACTTAAAGGATTAGTGTGAAGGTTAAAATTCCTAAAGTAACAA  
 GGATATTCAAAATCTGTGTGTAGCAGGTGAAGATAAGTAAAGAACAGCAATAGAGTTAAG  
 TAGCAAAACTATACTGTATGATGACAATGCAATAACTAAAGATCCAGTTAACATGTTG  
 TTAGTGGTCTCAAGCTGTAGTTGATGGGCTTGTAGCTTACCTTTAGCATCTGAAAGA  
 GCTCCCTTACTTAACTTACAGCAGGAAAATTAGATTCAGTGTAAAGCTGAGTTGAAAAGA  
 GTAATGATTAAAATCTACAAACAGGTGTAATACTTCTAAAAGGTTACTTAGCTGGTGA  
 GTAAAACCTATATCTAAAGATGTAGAAAATGATTAAAAGATATGGGACTTAAAGGTTAAG  
 ATTATCAGGAGATGATAGATGAAACTTCTTAGCTATAGCTGTAGAAATAGGTTGATAA  
 TGATAAACTGTTGTAGTGTGGAGGAACAGGATTAGCGGTAGCTATGAGTAACTCCAGTGC  
 TTCTCAATAAGAACATCAATGGAGAACACTGACTTAAAGGTGATGCAACTCCAAATGAGT  
 TGTGATGAAAAGCTAAAGTATAAATTCTGAAGTAAAGGTTAGTCTAGCATCACAAGT  
 TGATAATAGTGGTGTAAATAGTGTCTTCAAAGGTAATGGGAACAGATGATGTGCTAC  
 TGGAAAATCACCTGAGGAGATATGGAGAAGTAGACAGAACAAATGCTAAAGTATAA  
 AAGAAGATGATTCTTAAATGGAGAAGTTACAAACTCTTGTAGCTAAAGTGGTCAA  
 CTAAGAAGATCAATTAGTAGATGTCTTAGCAGGTGCTGCAATTGCTGGTAACTTGGTGTA  
 CAGTAGATAATGAGGAAAACCTACAGTGTGCTATAAAAAGCTTCCAGCACCAATTGTT  
 TAGCAACAGATTCTTATCTGTACAAATGTAGCTATAAGTAAAGCTGTAATGTGAGC  
 CTAATACTAAGAACTAGTTCAAGTGTGAAAGGTAGCTACTTCAGTGTGAAAGTAA  
 AAGATTATTAGATATG

## SEQ ID No. 9

(Strain 170444 DNA)

ATGAAATAAGAAAAATATAGCAATAGCTATGTCAGGTTAACAGTTTAGCTTCGCTGCTCT  
 GTTTTGTGCAACTACTGGAACACAAGGTTACTGTAGTTAAAACGACTGGAAAAAGCA  
 GTAAAAACAACTTAAAGATGGACTAAAGGATAATAGTATGGAAAGATAACTGTATCTTAA  
 TGATGGGGTTGTGGGTGAAGTAGCTCTTAAAGTGTCAATAAGAACGGGACAGAGATGCT  
 CAGCTGAGAAGTTATATAACTCTGAATTAGATAAATTAGGTGATGGAGGATTATG  
 TTGATTTTCTGTAGTTATAATTAGGAAAAAAAATAACTAATCAAGCAGATGCGAAG  
 CAATTGTTACAAAGTTAAATTCTTAATGAGAAAATTCTTATGATATAGCAACTAAAGATA  
 CTTTGGATGGTAGTAAACACAAGATAGTGAAGGTTAAAATGTTGCTGCAACAAAGGCA  
 CTTAAAGTTAAAGATGTGCTACATTGGTTGAAGTCTGGGAAAGCGAAGATAGTGGGAT  
 GTTATTGAAATGAAAGCAGGAGCTGTAGAGGATAAGTAGTGTAAAGTGGGAGATAGTACGGC  
 AGGTTATGCAATAATCTCTCTAGTACTGGACTTGTAAATGAGGTTAAAGGAAACAAATTG  
 TTAAATAAAACTTAAAGTTGTAGTACAGGTGTTCAACACCTAGTGTGTAAGTGTGAA  
 TGGTTTGTAACTTAAAGATGTACTGATTAGCAAAATCAGGTACTATAAATGTAAGAGTT  
 AAATGCAAAAAGGAAAGATCAATTGTATAGTGCAGCTCATACATCAGCTGAAATTG  
 CTAAAGACATGTTATTGCAAGTAAATTCTGTAAGCATAAAAGGCAATAGTAGCATTAC  
 AAAATGATGTTAGAGTCTAATTAGTTCTAGTTAGTAAATGGAAAATATCAAGTGT  
 ATCCAGAAGGTTAAAGGAGGTTATGAAACTAAATCAGCAATGATACAATAGCTAGTCAAGG  
 CCAGCTAAAGTAGTTATAAAAAGCTAAATAAATTAAAGGTTAAAAGGTTAGTGTAGT  
 AAAACATATAATAACTTATTCAAATGTTAACAGTAGCAGGAGAAGGATAAGTAAAGAAC  
 TGCTATAGAATAAGTAGAAATATTATAATTCTGTAGTGTGATAAAAATGCAATAACTGATAAAGC  
 AGTTAATGATATAGTATTAGTTGATCTACATCTATAGTTGATGGTCTGTGCTACCACTTA  
 GCTTCAGAAAAAACAGCTCAATTATTAACTTAAAGGATAAAATGTTCACTAGT  
 TCTGAAATAAGGAGGTTATGAACTTAAAGGAGCAGACTGGTATAAATACTCTAAAGGTT  
 TATTAGCTGGGGAGTTATTCTATATCTAAAGATGTAGAAAATGAAATTGAAAACATGGGT  
 CTTAAAGTTAGGATTAATCAGGAGAAGACAGATCAGAACCTTCTTAGCAATAAGCTGATGAA  
 ATAGGTCTGATAATGATAAAAGCATTTGAGTTGTTGGTGTAGCTGGGTTAGCAGATGCTATGAGT  
 ATAGCTCCAGTTGCTCTCAACTTAAAGATGGAGATGCTACTCCAAATAGTAGTTGATGTTG  
 AAAGCAGAAAAGGAAATAAGTGTGATGCTAAAGGTTCTTAGGAACTTCTGATGTTGATAATA  
 GGTGGAAAATAGCGTACTAAAGGAGATGAGGAGCTAAGTGTGCAACTGGTAAAGGTT  
 TCCAGATAGATAAGTGGAGATGATAGACAAGCAACTATGCTGAAAGTTAAAAGAAGATG

ATTATTTACAGATGGTGAAGTTGTAATTACTTTGTCAAAAGATGGTCTACTAAAGAAG  
ATCAATTAGTAGATGCCCTAGCAGCACCAATAGCAGGTAGATTTAAGGAGTCAGCTC  
CAATCATACTAGCTACTGATACTTATCTCTGACCAAAATGTAGCTGTAAAGCAGTTC  
CTAAAGATGGTGAACTAACITAGTCAAGTAGGAAAGGTATAGCTTCAGTTAAACAA  
AAATGAAAGATTATTAGATATGA

SEQ ID No. 10

(Strain 170426 DNA)

ATGAATAAGAAAAATATAGCAATAGCTATGTCAAGGTTAACAGTTAGCTCGCTGCTCCT  
GTTTTGCTGCAACTCTGGAACACAAGGGTATACTGTAGTTAAAACGACTGGAAAAAGCA  
GTAAACAAATTACAGGATGGACTAAAGATAATAGTATAGGAAAGATAACTGTATCTTAA  
TGATGGGGTTGTGGGTGAAGTAGCTCTAAAAGTGTAAATAAGAAAGCCGACAGAGATGCTG  
CAGCTGAGAAGTTATAACTCTGTAAACTCAATAGATAAATTAGGTGATGGAGATTATG  
TTGATTTCTGTAGTATTATAATTAGAAGAAAAAAATAATAACTAAGCAGATGCGAGAAG  
CAATTGTTACAAAGTTAAATTCACTTAATGAAAGAAACTCTTATTGTATAGCAACTAAAGATA  
CTTITGGATGGTTAGTAAACACAAGATAGTGAAGGTAAGGAAATAATGTGCTGCAACAAAGGC  
CTTAAAGTTAAAGATGGTCTACATTGGTTGAAGTCTGGTGAAGCGAAGATACTGGATAT  
GTGTTGAAATGAAAGCAGGAGCTGTAGAGGGATAATGTTAAAGTGGAGATAGTACGGC  
AGGTATTGCAATAAATCTCTCTAGTACTGGACTCTGAATATGCAAGGTGTTCAACACCTAGTGTGACTGTAA  
TTTAATAAAACTCTAAAGTTGATGTAACAGGTGTTCAACACCTAGTGTGACTGTAA  
TGGTTTGTAACTAAGATGACTGTATTGCAAATACTGGTACTATAATGTAAGGTT  
AAATGCAAAGAAGAATCAATTGATATAGTCAAGCTCATACATCAGCTGAAATTAG  
CTTAAAGATATGATTGATCCAGATGAAATCTGTAAGCATAAAGGCAATAGTAGCATTAC  
AAAATGATGTTAGAGTCTAATTGTTGAGTTAGTAAATGAAAGTCAAGTGTATT  
ATCCAGAAAGGTTAAAGGTTAGAAACTAACGCAATGATACATACTAGTGTCAAGATACA  
CCAGCTAAAGTAGTTAATAAAAGCTAATAAATGTTAAAGGTTAAAGGTTAGTGTGATT  
AAAACATTAATATACTTATCAACTAAGTGTAAACAGTCAGGAGAAGATGAAAC  
TGCTATAGAATTAAAGTAGTAAATTATAATCTGTATGATAAAATGCAATACTGATAAGC  
AGTTAATGATATGTTAGTTGGATCTACATCTAGTTGATGGCTTGTGCACTACCA  
GCTTCAGAAAAACAGCTCCATTATTAAACTCAAAGGATAAAATTAGATTCTCATGTA  
TCTGAAATAAAGAGCTTGTGAACTTAAAGAGTCACAGGTATAAATACTCTAAAAAGTT  
TATTAGCTGGTGGAGTTAAATTCTATCTAAAGATGTTGAGAAAATGAAATTGAAAAACATGGGT  
CTTAAGTTACTAGATTATCAGGAGAAAGACAGATACGAAACTCTCTAGTAACTGCTGAA  
ATAGGCTCTGATAATGATAAAGCATTGTTAGTTGGTGTACTGGATTAGCAGATGCTATGAGT  
ATAGCTCCAGTTGCTCTCAACTTAAAGATGAGGATGCTACTCCAAATAGTAGTTGAGTGG  
AAAGCAAAAGAAATAAGTGTGATGCTAAAGGTTCTAGGAACCTTGTATGTTGATATAATA  
GGTGGAAAAATAGCGTATCTAAAGAGATGTAAGAGTCATAAGATAGTGCAACTGGAAAAAC  
TCCAGATAGAATAAGTGGAGATGATAGACAAGCAACTAATGCTGAGTTTTAAAGAAGATG  
ATTATTCACAGATGGTGAAGTTGTAATTACTTGTGCAAAAGATGGTCTACTAAAGAAG  
ATCAATTAGTAGATGCCCTAGCAGCAGCACCAATAGCAGGTAGATTTAAGGAGTCAGCTC  
CAATCATACTAGCTACTGATACTTATCTCTGACCAAAATGTAGCTGTAAAGTAAAGCAGTTC  
CTAAAGATGGTGAACTAACITAGTCAAGTAGGAAAGGTATAGCTTCAGTTAAACAA  
AAATGAAAGATTATTAGATATG

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
15 August 2002 (15.08.2002)

PCT

(10) International Publication Number  
**WO 02/062379 A3**

(51) International Patent Classification<sup>7</sup>: **C12N 15/31.** C07K 14/33, C12N 15/62, C07K 16/12, A61K 39/08, 31/711

(21) International Application Number: **PCT/IE02/00017**

(22) International Filing Date: 11 February 2002 (11.02.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
2001/0137 9 February 2001 (09.02.2001) IE

(71) Applicant (*for all designated States except US*): **THE PROVOST, FELLOWS AND SCHOLARS OF THE COLLEGE OF THE HOLY AND UNDIVIDED TRINITY OF QUEEN ELIZABETH [IE/IE]**; Near Dublin, College Green, Dublin 2 (IE).

(74) Agent: **O'BRIEN JOHN A AND WELDON, Michael J.** c/o John A. O'Brien & Associates, Third Floor, Duncairn House, 14 Carysfort Avenue, Blackrock, County Dublin (IE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GII, GM, HR, IU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GII, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:  
16 January 2003

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **DOYLE, Rachael** [IE/IE]; 19 Deerpark Avenue, Castleknock, Dublin 15 (IE). **KELLEHER, Dermot** [IE/IE]; 30 Royal Terrace West, Dun Laoghaire, County Dublin (IE). **WINDLE, Henry, J.** [IE/IE]; 15 Cherryfield Avenue Upper, Ranelagh, Dublin 6 (IE). **WALSH, James, Bernard** [IE/IE]; 3 Ardlin Park, Blackrock, County Dublin (IE). **DEIRDRE, Ni, Eldhia** [IE/IE]; 15 Watkins Buildings, The Coombe, Dublin 8 (IE).

WO 02/062379 A3

(54) Title: CLOSTRIDIUM DIFFICILE VACCINE

(57) Abstract: A vaccine for the treatment or prophylaxis of *C. difficile* associated disease comprises a *C. difficile* gene or a *C. difficile* peptide/polypeptide or a derivative or fragment or mutant or variant thereof which is immunogenic in humans. The gene encodes a *C. difficile* surface layer protein, SlpA or variant or homologue thereof. The peptide/polypeptide is a *C. difficile* surface layer protein, SlpA or variant or homologue thereof. The vaccine may comprise a chimeric nucleic acid sequence.

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IE 02/00017

A. CLASSIFICATION OF SUBJECT MATTER  
 IPC 7 C12N15/31 C07K14/33 C12N15/62 C07K16/12 A61K39/08  
 A61K31/711

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
 IPC 7 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EP0-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, EMBASE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 20304 A (ORAVAX INC) 29 April 1999 (1999-04-29)  page 22 ---	1,2,5,7, 19, 23-37, 52-54, 57-64
P,X	CALABI EMANUELA ET AL: "Molecular characterization of the surface layer proteins from Clostridium difficile." MOLECULAR MICROBIOLOGY, vol. 40, no. 5, June 2001 (2001-06), pages 1187-1199, XP002946325 ISSN: 0950-382X Table 1: Strain 1, 33 kDa band ---	1-9,11, 19-21, 23-42, 44,52-64  -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the International filing date

"L" document which may throw doubts on priority, claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the International filing date or priority date and not in conflict with the application but which understanding the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step if it is made obvious by the document alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"a" document member of the same patent family

Date of the actual completion of the international search

16 August 2002

Date of mailing of the international search report

22.10.2002

Name and mailing address of the ISA

European Patent Office, P.B. 6018 Patentent 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, TX. 51 651 epo nl,  
Fax: (+31-70) 340-3015

Authorized officer

Mata-Vicente, M

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IE 02/00017

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>KARJALAINEN TUOMO ET AL: "Molecular and genomic analysis of genes encoding surface-anchored proteins from Clostridium difficile."  <b>INFECTION AND IMMUNITY,</b>  vol. 69, no. 5, May 2001 (2001-05), pages 3442-3446, XP002946326  ISSN: 0019-9567  Associated to Acc. N°: AJ291709.  -----</p>	1-9,11, 19-21, 23-42, 44,52-64
A	<p>CERQUETTI M ET AL: "CHARACTERIZATION OF SURFACE LAYER PROTEINS FROM DIFFERENT CLOSTRIDIUM DIFFICILE CLINICAL ISOLATES" MICROBIAL PATHOGENESIS, ACADEMIC PRESS LIMITED, NEW YORK, NY, US, vol. 28, no. 6, June 2000 (2000-06), pages 363-372, XP002946324  ISSN: 0882-4010  -----</p>	
A	<p>MASTRANTONIO P ET AL: "Identification of Clostridium difficile genes encoding surface proteins with adhesive properties."  ABSTRACTS OF THE GENERAL MEETING OF THE AMERICAN SOCIETY FOR, vol. 100, 2000, page 72 XP001002649  100th General Meeting of the American Society for Microbiology; Los Angeles, California, USA; May 21-25, 2000, 2000  ISSN: 1060-2011  the whole document  -----</p>	

**INTERNATIONAL SEARCH REPORT**International application No.  
PCT/IE 02/00017**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:  
**Although claims 57, 63 and 64 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.**
2.  Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
**see FURTHER INFORMATION sheet PCT/ISA/210**
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

**see additional sheet**

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
**(9, 11, 21, 42, 44) - (1-8, 19, 20, 23-41, 52-64) - partial**

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: (9, 11, 21, 42, 44) - complete; (1-8, 19, 20, 23-41, 52-64) - partial

Clostridium difficile S layer protein (SipA) comprising SEQ ID NO:1 and its corresponding gene (*sipA*), which comprises SEQ ID NO:3; epitopes, homologs, derivatives, variants or fragments thereof. Chimeras comprising any of the previously mentioned polynucleotides/(poly)peptides. Antibodies against those (poly)peptides. Vaccines comprising any of the former and methods for prophylaxis/treatment of C. difficile-associated diseases based on the use thereof.

2. Claims: (10, 13, 14, 17, 18, 22, 43, 46, 47, 50, 51) - complete; (1-8, 19, 20, 23-41, 52-64) - partial

Idem as subject 1, but restricted to a gene comprising SEQ ID NOS: 5, 6, 9 or 10 and a polypeptide/peptide comprising SEQ ID NO:2.

3. Claims: (12, 45) - complete; (1-8, 19, 20, 23-37, 39-41, 52-54, 56-64) - partial

Idem as subject 1, but restricted to a gene comprising SEQ ID NO:4.

4. Claims: (15, 48) - complete; (1-8, 19, 20, 23-37, 39-41, 52-54, 56-64) - partial

Idem as subject 1, but restricted to a gene comprising SEQ ID NO:7.

5. Claims: (16, 49) - complete; (1-8, 19, 20, 23-37, 39-41, 52-54, 56-64) - partial

Idem as subject 1, but restricted to a gene comprising SEQ ID NO:8.

6. Claims: (65) - complete

Use of interleukin 12 as an adjuvant in a C. difficile vaccine.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

**7. Claims: (66) - partial**

Use of humanised antibodies for passive vaccination of an individual with C. difficile infection.

**8. Claims: (66) - partial**

Use of serum for passive vaccination of an individual with C. difficile infection.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Please notice that:

1. The translations of the ORFs contained in appendices 1-8 are not included in the sequence listing and, therefore, the one corresponding to the first invention has not been searched. In case the applicant decided to pay additional fees, he should be aware of the fact that the same will apply to the other inventions..
2. Claims 39 and 56 refer to SEQ ID NOS:3-10 as "amino acid sequences" but, actually, they are nucleotidic sequences.
3. The sequence numbering is confusing. The sequence identity numbers mentioned in the description and claims do not correspond with those of the sequence listing (example: SEQ ID NO:1 of the description is a peptide which appears under SEQ ID NO:9 of the sequence listing).

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

## INTERNATIONAL SEARCH REPORT

## Information on patent family members

International Application No

PCT/IE 02/00017

Patent document cited in search report	Publication date		Patent family member(s)	Publication date
WO 9920304	A 29-04-1999	AU CA EP WO US	1108299 A 2307331 A1 1024826 A1 9920304 A1 6214341 B1 2001051153 A1	10-05-1999 29-04-1999 09-08-2000 29-04-1999 10-04-2001 13-12-2001